PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/52, 15/10, 9/92, 9/10, 1/21, C12Q 1/68

A1

(11) International Publication Number: **WO 97/08323**

(43) International Publication Date: 6 March 1997 (06.03.97)

(21) International Application Number: PCT/EP96/03643

(22) International Filing Date: 19 August 1996 (19.08.96)

(30) Priority Data:
95810534.8 30 August 1995 (30.08.95) EP

(34) Countries for which the regional or international application was filed: DE et al.

(71) Applicant *(for all designated States except US)*: CIBA-GEIGY AG [CH/CH]; Klybeckstrasse 141, CH-4002 Basle (CH).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: SCHUPP, Thomas [CH/CH]; Fröschmattweg 5, CH-4313 Möhlin (CH). ENGEL, Nathalie [FR/FR]; 29, rue de la Doller, F-68260 Kingersheim (FR). BIETENHADER, Jürg [CH/CH]; Röserenstrasse 17, CH-4410 Liestal (CH). TOUPET, Christiane [FR/FR]; 12/5, rue de l'Ours, F-68200 Mulhouse (FR). POSPIECH, Andreas [DE/DE]; Scheffelstrasse 23, D-79102 Freiburg (DE).

(74) Common Representative: NOVARTIS AG; Patent- und Markenabteilung, Klybeckstrasse 141, CH-4002 Basle (CH).

(81) Designated States: AL, AU, BB, BG, BR, CA, CN, CU, CZ, EE, GE, HU, IL, IS, JP, KP, KR, LK, LR, LT, LV, MG, MK, MN, MX, NO, NZ, PL, RO, SG, SI, SK, TR, TT, UA, US, UZ, VN, ARIPO patent (KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(54) Title: STAUROSPORIN BIOSYNTHESIS GENE CLUSTERS

(57) Abstract

The present invention relates especially to a DNA fragment that is obtainable from the gene cluster within the genome of *Streptomyces* or *Actinomyces* that is responsible for staurosporin biosynthesis and that contains at least one gene or a part of a gene that codes for a polypeptide that is involved directly or indirectly in the biosynthesis of staurosporin and to methods of preparing said DNA fragment. The present invention relates furthermore to recombinant DNA molcules containing one of the DNA fragments according to the invention and to the plasmids and vectors derived therefrom. Also included are host organisms transformed with the said plasmid or vector DNA.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	**GB**	United Kingdom	**MW**	Malawi
AT	Austria	**GE**	Georgia	**MX**	Mexico
AU	Australia	**GN**	Guinea	**NE**	Niger
BB	Barbados	**GR**	Greece	**NL**	Netherlands
BE	Belgium	**HU**	Hungary	**NO**	Norway
BF	Burkina Faso	**IE**	Ireland	**NZ**	New Zealand
BG	Bulgaria	**IT**	Italy	**PL**	Poland
BJ	Benin	**JP**	Japan	**PT**	Portugal
BR	Brazil	**KE**	Kenya	**RO**	Romania
BY	Belarus	**KG**	Kyrgystan	**RU**	Russian Federation
CA	Canada	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CF	Central African Republic			**SE**	Sweden
CG	Congo	**KR**	Republic of Korea	**SG**	Singapore
CH	Switzerland	**KZ**	Kazakhstan	**SI**	Slovenia
CI	Côte d'Ivoire	**LI**	Liechtenstein	**SK**	Slovakia
CM	Cameroon	**LK**	Sri Lanka	**SN**	Senegal
CN	China	**LR**	Liberia	**SZ**	Swaziland
CS	Czechoslovakia	**LT**	Lithuania	**TD**	Chad
CZ	Czech Republic	**LU**	Luxembourg	**TG**	Togo
DE	Germany	**LV**	Latvia	**TJ**	Tajikistan
DK	Denmark	**MC**	Monaco	**TT**	Trinidad and Tobago
EE	Estonia	**MD**	Republic of Moldova	**UA**	Ukraine
ES	Spain	**MG**	Madagascar	**UG**	Uganda
FI	Finland	**ML**	Mali	**US**	United States of America
FR	France	**MN**	Mongolia	**UZ**	Uzbekistan
GA	Gabon	**MR**	Mauritania	**VN**	Viet Nam

STAUROSPORIN BIOSYNTHESIS GENE CLUSTERS

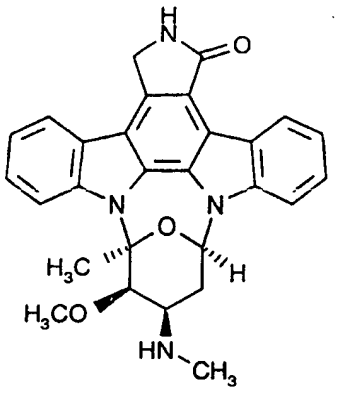
Staurosporin, an indole-carbazole alkaloid antibiotic, was first isolated from cultures of the microorganism *Streptomyces staurosporens* and described by Omura *et al.* (Omura *et al.*, J. Antibiot. (1977), **30**, 275-282). The biological properties of that secondary metabolite are of exceptional interest and include the following activities:

- inhibitory activity against fungi and yeasts (Omura *et al.*, J. Antibiot. (1977), **30**, 275-282),
- strong inhibition of Ca^{2+}/phospholipid-dependent serine/threonine protein kinases (PKC) (Tamoki *et al.*, Biochem. Biophys. Res. Comm. (1986), **135**, 397-402),
- antiproliferative activity (Tamoki *et al.*, Biochem. Biophys. Res. Comm. (1986), **135**, 397-402),
- inhibition of platelet aggregation (Oka *et al.*, Biol. Chem. (1986), **50**, 2723-2727).

The isoenzyme family of the protein kinase Cs (PKC) plays an important part in signal transduction and cell regulation (Nishizuka, Nature (1988), **334**, 661-665). The observation that phorbol esters, which have a tumour-stimulating property, stimulate PKC activity in cells (Nishizuka, Nature (1984), **308**, 693-698) led to the conclusion that the inhibition of those enzymes by staurosporin and by similar staurosporin-like compounds could perhaps be used in the chemotherapy of tumours.

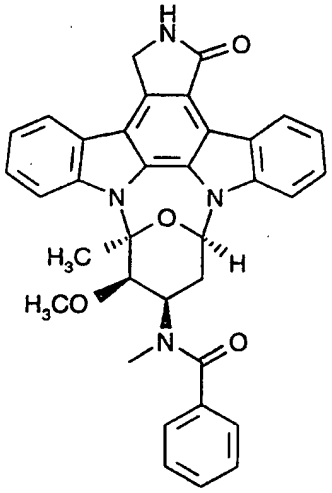
Later, staurosporins were isolated from other strains of Streptomyces, for example *Streptomyces longisporoflavus* (strain R-19, DSM 10189), *Streptomyces actuosus* (Morioka *et al.*, Agric. Biol. Chem. (1985), **49**, 1959-1963) and *Streptomyces* species, strain M-193 (Oka *et al.*, Biol. Chem. (1986), **50**, 2723-2727) and *Streptomyces* species, strain 383. Other alkaloids very similar to staurosporin, which contain the same chromophore as staurosporin and exhibit similar biological activity, have also been isolated. Examples are rebeccamycin (Nettleton *et al.*, Tetrahedron Lett. (1985), **26**, 4011-4014), UCN-01, UCN02 (Takahashi *et al.*, J. Antibiot. (1987), **40**, 1782-1783; Takahashi *et al.*, J. Antibiot. (1989), **42**, 571-576) and K-252 (Kase *et al.*, J. Antibiot. (1986), **39**, 1059-1065), which have also been described as PKC inhibitors or anti-tumour compounds.

Staurosporin has the structure of formula (1)



(1)

and is an exceptionally strong inhibitor of protein kinase C, but the molecule lacks the selectivity required for pharmaceutical applications involving the very specific inhibition of individual protein kinases. For that reason, analogous compounds based on the fermentation product staurosporin have been prepared by chemical derivatisation at different centres (Ruegg & Burgess, Trends in Pharmacological Science (1989), **10**, 218- 220). An example thereof is the compound of formula (2) (Meyer *et al.*, Int. J. Cancer (1989), **43**, 851-856)



(2)

which has selectivity for protein kinase C inhibition and exhibits antiproliferative activity *in vitro* and anti-tumour properties *in vivo*.

Streptomyces are gram-positive filamentous bacteria that are found ubiquitously in soil. *Streptomyces* cultures grow in the form of branching mycelia which, when nutrients are limited, are capable of differentiating further to form aerial mycelia and, finally, to form spores. A special property of that group of microorganisms is their enormous potential for producing an extremely large variety of differently structured metabolites, known as secondary metabolites. Many of those compounds have antibacterial, antifungal, anti-

tumour, immunomodulating or herbicidal properties and are therefore of great practical importance for pharmaceutical or agrochemical use.

Because of the practical importance of microbial secondary metabolites, there is a great deal of interest in understanding the genetic basis of their synthesis in order to create the means to influence them in a targeted manner. That is desirable especially because natural production strains, as in the case of the biosynthesis of staurosporin, generally yield only low concentrations of the secondary metabolites that are of interest. Those concentrations are not sufficient to satisfy the demand for the substance for wide-ranging activity tests and for preclinical and clinical trials, let alone for commercial production.

The genetic basis of secondary metabolite biosynthesis consists essentially in the genes that code for the individual biosynthesis enzymes and in the regulatory elements that control the expression of the biosynthesis genes. In all of the systems investigated hitherto, the secondary metabolite synthesis genes of *Streptomyces* have been found as clusters of adjacent genes. The size of such antibiotic gene clusters ranges from approximately 10 kilobases (kb) to approaching 100 kb. The clusters normally also contain specific regulator genes and genes for the resistance of the producing organism to its own antibiotic (Chater, Ciba Found. Symp. (1992), **171**, 144-162).

In the invention described herein, success has now been achieved, by identifying and cloning genes of staurosporin biosynthesis, in providing the genetic basis for improving in a targeted manner the productivity of staurosporin-synthesising *Streptomyces* and, especially, of *S. longisporoflavus* or, using genetic methods, for synthesising staurosporin analogues, such as other indole-carbazole alkaloids. In a first step, a staurosporin biosynthesis gene of *S. longisporoflavus* was successfully identified by complementation of a mutant blocked in a biosynthesis step and cloned. Using DNA sequencing, the expected function of the protein derived from the cloned gene in the relevant biosynthesis step of staurosporin was confirmed. On the basis of the DNA sequence, there was found on a cloned 2.1 kb BglII fragment a second gene that is involved in the synthesis of staurosporin and is likewise capable of complementing a mutant that is blocked in the synthesis of the sugar moiety of the staurosporin molecule. Finally, the cloned DNA fragment was used as a DNA probe for isolating the other staurosporin synthesis genes on large chromosomal DNA fragments of *S. longisporoflavus*.

The gene cluster thus isolated and characterised forms the basis for the targeted optimisation of staurosporin production in *S. longisporoflavus* and other *Streptomyces* or *Actinomyces*. The following molecular genetic objectives and/or techniques are of primary importance therein:

- overexpression of individual genes in production strains using plasmid vectors or by the incorporation of additional copies into the chromosome
- study of the expression and transcriptional regulation of the gene cluster during fermentation in different production strains and optimisation thereof by means of physiological parameters and appropriate fermentation conditions
- identification of regulator genes and of the DNA binding sites of the corresponding regulator proteins in the gene cluster. Characterisation of the effect of those regulatory elements on staurosporin production and influencing thereof by means of controlled mutations in those genes or in the DNA binding sites
- duplication of the whole gene cluster or of parts thereof in production strains.

In addition to its use for improving fermentative staurosporin production in accordance with the above description, the gene cluster can likewise be used for the biosynthetic preparation of novel staurosporin analogues. The following possibilities may be mentioned:

- inactivation of individual biosynthesis steps by means of gene disruption
- use of genes of the cluster as DNA probe for isolating from nature *Actinomyces* or other microorganisms that produce metabolites similar to staurosporin
- replacement of individual elements of the staurosporin gene cluster with those of other indole-carbazole alkaloid-producing *Actinomyces*, such as rebeccamycin, UCN-01, UCN-02 or K-252, and expression of novel, so-called hybrid metabolites.

Detailed description of the invention

The present invention relates to an isolated DNA fragment comprising a DNA region that is involved directly or indirectly in the biosynthesis of indole-carbazole alkaloids, including the adjacent DNA regions to the right and left which, because of their function in connection with indole-carbazole alkaloid biosynthesis, qualify as constituents of the indole-carbazole alkaloid gene cluster; and functional fragments thereof.

The present invention relates especially to an isolated DNA fragment comprising a DNA region that is involved directly or indirectly in the biosynthesis of staurosporin, including the

adjacent DNA regions to the right and to the left which, because of their function in connection with staurosporin biosynthesis, qualify as constituents of the staurosporin gene cluster.

The DNA fragments according to the invention may contain regulatory sequences, such as promotors, repressor or activator binding sites, repressor or activator genes or terminators; structural genes or information for enzymatic active domains. The invention relates also to any desired combinations of those DNA fragments with one another or with other DNA fragments, such as combinations of promotors, repressor or activator binding sites and/or repressor or activator genes from the indole-carbazole alkaloid gene cluster, especially the staurosporin gene cluster, with foreign structural genes, or combinations of structural genes from the indole-carbazole alkaloid gene cluster, especially the staurosporin gene cluster, with foreign promotors; and combinations of structural genes from different indole-carbazole alkaloid biosynthesis systems. Foreign structural genes code, for example, for proteins that are involved in the biosynthesis of other indole-carbazole alkaloids.

Preference is given to a DNA fragment comprising a DNA region that is involved directly or indirectly in the biosynthesis of staurosporin.

The DNA region or gene cluster described above contains, for example, the genes that code for the individual enzymes that are involved in the biosynthesis of the indole-carbazole alkaloids and especially of staurosporin, and the regulatory elements that control the expression of the biosynthesis genes. The size of such antibiotic gene clusters ranges from approximately 10 kilobases (kb) to approaching 100 kb. The gene clusters normally also contain specific regulator genes and genes for the resistance of the producing organism to its own antibiotic. There are to be understood as enzymes that are involved in the biosynthesis, for example, those that, starting from precursors of tryptophan and glucose, are required for the synthesis of the indole-carbazole alkaloids, such as staurosporin, such as methyl transferases, glucose epimerases, dTDP-glucose synthases (dTDP-glucose pyrophosphorylases), dCDP-glucose synthases (CTP-glucose synthases), hexose-1-P-nucleotidyl transferases, NDP-glucose 4,6-dehydratases, NDP-4-keto-6-deoxyhexose 3,5-epimerases, secondary metabolitic amino transferases, and enzymes for the conversion of l-tryptophan (2-molecules) into the indole-carbazole nucleus of staurosporin.

In a further preferred form, the DNA fragment according to the invention is obtained from the gene cluster within the genome of *Streptomyces* or *Actinomyces*, and especially of *Streptomyces longisporoflavus*, that is responsible for staurosporin biosynthesis.

For example, a DNA fragment according to the invention comprises a 35 kb DNA region as shown in Figure 2, and is preferably a DNA fragment that comprises a 10 kb region as shown in Figure 1. Special preference is given to a DNA fragment that contains one or more of the partial nucleotide sequences set out in SEQ ID NOs 1, 4 and 5, or functional fragments thereof, and any further DNA sequences in the vicinity of that sequence that, on the basis of homologies present, may be regarded as structural or functional equivalents and are therefore capable of hybridising with that sequence. Examples of other preferred DNA fragments are those that are obtainable in accordance with the method of the invention from the *Streptomyces longisporoflavus* genome and that overlap with the 2.1 kb fragment, such as the following fragments (see also Fig. 1):

- EcoRI: > 20 kb,
- PvuII: 3.5 kb and 6.5 kb;
- PvuI: 3.6 kb and 2.1 kb;
- BclI: 3.6 kb.

The DNA fragments according to the invention contain, for example, portions of sequence having homologies to the methyl transferases, to amino transferase or to enzymes that are involved in the synthesis of the deoxy sugar moiety of metabolites. In a preferred form, the DNA fragments according to the invention contain portions of sequence having homologies to the methyl transferases and the amino transferases of *Streptomyces* or *Actinomyces*, or glucose epimerases, such as dTDP-4-keto-6-deoxyglucose 3,5-epimerase; the DNA fragment according to the invention containing in an especially preferred form portions of sequence that code for a methyl transferase. Other especially preferred DNA fragments code for the proteins set out in SEQ ID NO 2 or SEQ ID NO 3, for the proteins represented by the open reading frames in SEQ ID NO 4, or for functional derivatives thereof in each case.

Preference is given also to DNA fragments containing portions of sequence that have homologies to the above-defined 35 kb DNA region or 10 kb DNA region or to SEQ ID NOs 1, 4 and 5 and that can therefore be used as a hybridisation probe within a genomic gene bank of an indole-carbazole alkaloid-producing organism, such as a staurosporin-producing organism, for detecting a constituent of the corresponding gene cluster. The DNA fragment

may comprise, for example, exclusively genomic DNA. Special preference is given to a DNA fragment containing the partial nucleotide sequence set out in SEQ ID NO 1, 4 or 5, or a sequence that, on the basis of homologies present, can be regarded as a structural or functional equivalent of the said partial sequence and is therefore capable of hybridising with that sequence.

In order to produce unambiguous signals during hybridisation, the DNA, bonded to filters (e.g. of nylon or nitrocellulose), is usually washed at 55-65°C in 0.2 × SSC (1 × SSC = 0.15 M sodium chloride, 15 mM sodium citrate).

The expressions 'homologies' and 'structural and/or functional equivalents' refer especially to DNA and amino acid sequences having few or minimal differences between the relevant sequences. Those differences can have very different causes. They may, for example, be mutations or strain-specific differences that occur naturally or are artificially induced or, alternatively, the observable differences with respect to the starting sequence are due to a specific modification that can be introduced, for example, as part of a chemical synthesis.

Functional differences can be regarded as minimal if, for example, the nucleotide sequence coding for a polypetide or a protein sequence has essentially the same characteristic properties as the starting sequence, whether it be in the area of enzymatic activity, immunological reactivity or, in the case of a nucleotide sequence, gene regulation.

Structural differences can be regarded as minimal provided that there is significant overlapping or similarity between the different sequences or that those sequences have at least similar physical properties. The latter include, for example, electrophoretic mobility, chromatographic similarities, sedimentation coefficients, spectrophotometric properties, etc..

In the case of nucleotide sequences, there should be at least 70 % identity, preferably 80 % and especially 90 % or more. In the case of the amino acid sequence, the corresponding values are at least 50 %, preferably 60 % and especially 70 %. An identity of 90 % is very especially preferred.

The invention relates also to a hybrid vector containing at least one DNA fragment according to the invention, such as a promotor, a repressor or activator binding site, a repressor or activator gene, a structural gene, a terminator or a functional moiety thereof. The hybrid vector contains, for example, an expression cassette containing a DNA fragment

according to the invention that is capable of expressing one or more proteins involved in indole-carbazole alkaloid biosynthesis, and especially in the biosynthesis of staurosporin, or a functional fragment thereof. The invention relates also to a host organism containing the hybrid vector described above.

Suitable vectors that form the starting point for the hybrid vectors according to the invention are generally known, such as pIJ702, pIJ486, pIJ487 and pIJ943.

Suitable host organisms within the scope of the invention are, for example, prokaryotic cells, such as *Actinomyces*, *Pseudomonades*, *E. coli*, or eukaryotic cells, such as yeasts and filamentous fungi. Examples of especially suitable host organisms are *Streptomyces*, such as *Streptomyces staurosporens*, *Streptomyces longisporoflavus*, *Streptomyces actuosus*, *Streptomyces* species, strain M-193 and *Streptomyces* species, strain 383.

The host organism can be transformed using generally customary methods, for example by means of protoplasting, Ca^{2+}, electroporation, viruses, lipid vesicles or a particle gun. The DNA fragments according to the invention may then either be present in the host organism as extrachromosomal constituents or may have been integrated into the chromosome of the host organism *via* suitable sections of sequence.

The invention relates also to a method of identifying, isolating and cloning a DNA fragment that is obtainable from the gene cluster within the genome of *Streptomyces* or *Actinomyces* that is responsible for indole-carbazole alkaloid biosynthesis, especially staurosporin biosynthesis, and that contains at least one gene that is involved directly or indirectly in the biosynthesis of indole-carbazole alkaloids, such as staurosporin, which method comprises the following steps:

a) constructing a representative gene library of an indole-carbazole alkaloid-producing organism, especially a staurosporin-producing organism, from the group of the *Streptomyces* or *Actinomyces*, which library contains substantially the entire genome divided into individual clones,
b) screening the said clones using a specific DNA probe that hybridises at least with a portion of the gene cluster responsible for the indole-carbazole alkaloid biosynthesis,
c) selecting the clones that allow a hybridisation signal with the DNA probe to be recognised; and
d) isolating a DNA fragment from the said clone that contains at least one gene that is involved directly or indirectly in the biosynthesis of the indole-carbazole alkaloid.

In a preferred form, the said staurosporin-producing organism is *Streptomyces staurosporens*, *Streptomyces longisporoflavus*, *Streptomyces actuosus*, *Streptomyces* species, strain M-193 or *Streptomyces* species, strain 383 or, especially, *Streptomyces longisporoflavus*.

The hybridisation probes used are, for example, one of the DNA fragments according to the invention. There may also be used as hybridisation probe sections of sequence originating from the right- and/or left-hand margins of the said DNA fragments.

Special preference is given to a method of identifying and isolating all of the DNA sequences that are involved in the construction of an indole-carbazole alkaloid gene cluster, which method comprises:

a) constructing a representative gene library of an indole-carbazole alkaloid-producing organism from the group of the *Streptomyces* or *Actinomyces*, which library contains substantially the entire bacterial genome divided into individual clones;
b) hybridising the said clones, using as probe molecule one of the previously isolated DNA fragments or selected portions thereof that overlap at least with a portion of the adjacent DNA regions to the right and/or left within the gene cluster;
c) selecting the clones that allow a strong hybridisation signal with the DNA probe to be recognised;
d) isolating the fragments containing overlapping DNA regions from the clones selected in accordance with (c) and isolating the fragment that projects furthest beyond the overlapping region;
e) testing the DNA fragment isolated in accordance with (d) for its ability to function within the gene cluster;
f) if it can be demonstrated that the DNA fragment isolated in accordance with (d) functions in the context of the indole-carbazole alkaloid biosynthesis, repeating the method according to steps (a) to (e), the DNA fragment isolated in accordance with (d), or selected portions thereof, especially those from the left- and/or right-hand margin of the said fragment, now acting as the DNA probe, until in the function test for each newly isolated DNA fragment no further functioning is detected in the context of the indole-carbazole alkaloid biosynthesis and the end of the gene cluster has thus been reached; and
g) carrying out the method according to steps (a) to (f), if necessary in the other, not hitherto selected, direction.

In order to isolate the DNA fragments according to the invention, the genomic gene banks that synthesise the desired indole-carbazole alkaloid, especially staurosporin, are first produced from the organism strains of interest.

Genomic DNA can be obtained from a host organism in a variety of ways, for example by extraction from the nuclear fraction and purification of the extracted DNA by known methods.

The fragmentation of the genomic DNA to be cloned to a size suitable for insertion into a cloning vector, which fragmentation is required for the production of a representative gene bank, can be effected either by mechanical cutting or, preferably, by cleavage with suitable restriction enzymes. Special preference is given within the scope of this invention to partial cleavage of the genomic DNA, leading to overlapping DNA fragments.

Suitable cloning vectors, which are already used routinely for the production of genomic gene libraries, include, for example, cosmid vectors, plasmid vectors or phage vectors.

Suitable clones containing the desired gene(s) or gene fragment(s) can then be obtained from the gene libraries produced in that manner, using a screening programme.

One possible method of identifying the desired DNA region is, for example, to transform strains that, because of a blocked synthesis path, are not capable of producing staurosporin or other indole-carbazole alkaloids, using the gene bank described above, and to identify those clones which after the transformation are again capable of producing staurosporin (revertants). The vectors that lead to the revertants contain a DNA fragment required in staurosporin synthesis.

A further possible method of identifying the desired DNA region is based, for example, on the use of suitable probe molecules (DNA probe) which are obtained, for example, as described above. Various standard methods are available for identifying suitable clones, such as differential colony hybridisation or plaque hybridisation. When expression gene banks are used, it is possible, moreover, to use immunological detection methods based on the identification of specific translation products.

There may be used as probe molecule, for example, a previously isolated DNA fragment from the same gene or from a structurally related gene that, because of the homologies that are present, is capable of hybridising with the corresponding section of sequence within the desired gene or gene cluster to be identified. Preference is given within the scope of the present invention to the use as probe molecule of a DNA fragment obtainable from a gene or another DNA sequence that plays a role in the synthesis of staurosporin.

If the amino acid sequence of the gene to be isolated, or at least parts of that sequence, are known, it is possible on the basis of that sequence information, in an alternative form of the method, to use an appropriate synthesised DNA sequence for the hybridisations or PCR amplifications.

In order to make the desired gene or parts of a desired gene easier to detect, one of the DNA probe molecules described hereinbefore can be labelled with a suitable readily detectable group. There is to be understood by 'detectable group' within the context of this invention any material that has a specific easily identifiable physical or chemical property.

Special mention may be made at this point of enzymatically active groupings, such as enzymes, enzyme substrates, coenzymes and enzyme inhibitors, also fluorescent and luminescent agents, chromophores and radioisotopes, such as ^{3}H, ^{35}S, ^{32}P, ^{125}I and ^{14}C. The ready detectability of those labels derives on the one hand from their inherent physical properties (e.g. fluorescent labels, chromophores, radioisotopes), and on the other hand from their reaction and binding properties (e.g. enzymes, substrates, coenzymes, inhibitors). Such materials are already widely used, especially in the area of immunoassays, and in the majority of cases can also be used in the present Application.

General methods relating to DNA hybridisation are described, for example, in Maniatis T. *et al.* (1982).

Those clones within the gene libraries described hereinbefore that are capable of hybridising with a probe molecule and that can be identified using one of the detection methods mentioned above can then be analysed further in order to determine in detail the extent and the nature of the coding sequence.

An alternative method of identifying cloned genes is based on the construction of a gene library made up of plasmid or expression vectors. In that method, analogously to the

methods already described hereinbefore, genomic DNA containing the desired gene product is first isolated and then cloned into a suitable plasmid or expression vector. The gene libraries thus produced can then be screened by suitable methods, for example using complementation studies, and the clones that contain the desired gene or at least a portion of that gene as an insert can be selected.

Using the methods described hereinbefore, it is thus possible to isolate a gene that codes for a specific gene product.

For the purpose of further characterisation, the DNA sequences purified and isolated in the manner described hereinbefore are subjected to restriction analysis and to sequence analysis.

For sequence analysis, the previously isolated DNA fragments are first cut into fragments with the aid of suitable restriction enzymes and then cloned into suitable cloning vectors. In order to avoid sequencing errors, it is advantageous to sequence both DNA strands completely.

Various alternative methods are available for analysing the cloned DNA fragment in respect of its function in the context of staurosporin biosynthesis.

For example, it is possible using complementation experiments with defective mutants not only to establish that a gene or gene fragment is in principle involved in the biosynthesis of secondary metabolites, but in addition to verify the specific synthesis step in which the said DNA fragment is involved.

In an alternative form of analysis, the evidence is obtained in exactly the opposite way. By transferring plasmids containing DNA sections having homologies to corresponding sections on the genome, the said homologous DNA sections are integrated *via* homologous recombination. If, as in the present case, the homologous DNA section is a region within an open reading frame of the gene cluster, the plasmid integration leads to inactivation of the gene as a result of gene disruption and, consequently, to interruption of the production of secondary metabolites. On the basis of current knowledge, it is assumed that a homologous region comprising at least 100 bp, and preferably more than 1000 bp, is sufficient to bring about the desired recombination event.

Preference is given, however, to a homologous region extending over a range of from 0.3 to 4 Kb, especially over a range of from 1 to 3 Kb.

For the production of suitable plasmids having sufficient homology for integration *via* homologous recombination, a subcloning step is preferably provided in which the previously isolated DNA is digested and fragments of suitable size are isolated and then cloned into a suitable plasmid. Suitable plasmids are, for example, the plasmids generally used for genetic manipulations in *Streptomyces*, such as pIJ486, pIJ487 and pGMI60.

In principle, it is possible to use any current cloning vectors for the production and replication of the constructs described hereinbefore, for example plasmid or bacteriophage vectors, provided that they have replication and control sequences originating from species compatible with the host cell.

As a rule, a cloning vector carries a replication origin and also specific genes that lead to phenotypic selection features in the transformed host cell, especially resistance to antibiotics. The transformed vectors can be selected on the basis of those phenotypic markers after transformation in a host cell.

Selectable phenotypic markers that can be used within the context of this invention include, for example, without this representing a limitation of the subject of the invention, resistance to thiostreptone, ampicillin, tetracycline, chloramphenicol, hygromycin, G418, kanamycin, neomycin or bleomycin. Prototrophy for specific amino acids can, for example, act as a further selectable marker.

Preference is given within the scope of the present invention especially to *Streptomyces* and *E. coli* plasmids, such as the plasmids puC18, pUC19 and pIJ486 used in the present invention.

Suitable host cells for the cloning described hereinbefore are, according to this invention, especially prokaryotes, including bacterial hosts, such as *Streptomyces*, *Actinomyces*, *Pseudomonades* or salmonella.

Special preference is given to *E. coli* hosts, such as the *E. coli* strain HB101 or X-1 Blue MR ® (Stratagene), or *Streptomyces*, such as strain TK23.

Competent cells of the *E. coli* strain HB101 are produced by the methods customarily used for the transformation of *E. coli*. For *Streptomyces* the transformation method according to Hopwood *et al.* (Genetic manipulation of *Streptomyces* a laboratory manual. The John Innes Foundation, Norwich (1985)) is customarily used.

After transformation and subsequent incubation on a suitable medium, the resulting colonies are subjected to differential screening by plating out onto selective media. The corresponding plasmid DNA can then be isolated from the colonies containing plasmids having cloned-in DNA fragments.

A DNA fragment according to the invention that contains a DNA region involved directly or indirectly in the biosynthesis of staurosporin and that is obtainable in the manner described hereinbefore from the gene cluster of the staurosporin biosynthesis can also be used as a starting clone for the identification and isolation of other, adjacent DNA regions from the said gene cluster that overlap therewith.

That can be achieved, for example, within a gene library consisting of DNA fragments having overlapping DNA regions, by means of 'chromosome walking' using the previously isolated DNA fragment or, especially, its 5' or 3' end sequences. The procedures for chromosome walking are known to a person skilled in the art. Details can be obtained, for example, from the publications of Smith *et al.* (Methods Enzymol (1987), **151**, 461-489) and Wahl *et al.* (Proc Natl. Acad. Sci, USA (1987), **84**, 2160-2164).

A precondition for chromosome walking is the presence within a gene library of clones having DNA fragments that are as long and cohesive as possible and that overlap one another to the greatest possible extent, and of a suitable starting clone that contains a fragment located in the vicinity of or, preferably, inside the region to be analysed. If the precise location of the starting clone is unknown, the walking is preferably carried out in both directions.

The actual walking step begins by using the starting clone, once identified and isolated, as a probe in one of the hybridisation reactions described hereinbefore to trace adjacent clones, which have regions that overlap with the starting clone. By means of hybridisation analysis, the fragment that projects furthest beyond the overlapping region can be determined. That fragment is then used as the starting clone for the second walking step, there being determined in this case the fragment that overlaps with the said second clone in

the same direction. In that manner, by means of continuous walking forward along the chromosome, a collection of overlapping DNA clones covering a large DNA region is obtained. Those clones can then be ligated together by known methods, if necessary after carrying out one or more subcloning steps, to form a fragment comprising some or, preferably, all of the components essential for staurosporin biosynthesis.

In the hybridisation reaction for identifying clones having overlapping margins, preference is given to the use of a part fragment from the left- or right-hand margin, which can be obtained by means of a subcloning step, instead of a very large and unwieldy whole fragment. Because of the relatively small size of the said part fragment, fewer positive hybridisation signals are obtained in the hybridisation reaction, with the result that the analysis requires markedly less effort than when the whole fragment is used. It is also advisable for the part fragment to be characterised in detail in order to exclude the possibility that it contains relatively large amounts of repetitive sequences, possibly scattered over the entire genome, which would make a target-specific walking step sequence very much more difficult.

Since the gene cluster responsible for staurosporin biosynthesis covers a relatively large region of the genome, 'large-step walking' or cosmid walking is advantageous according to the present invention. Using cosmid vectors, which allow the cloning of very large DNA fragments, it is possible in those cases to cover a very large DNA region, which may comprise up to 45 Kb, in a single walking step.

In one form of the present invention, for example, for the construction of a cosmid gene bank of *Streptomyces* or *Actinomyces*, total DNA of the order of magnitude of DNA fragments of approx. 100 kb is isolated and then partially digested with the aid of suitable restriction endonucleases.

The digested DNA is then extracted in customary manner in order to remove any remaining endonucleases, precipitated and, finally, concentrated. The resulting fragment concentrate is then separated, for example by means of density gradient centrifugation, according to the size of the individual fragments. When the fractions thus obtainable have undergone dialysis, they can be analysed on an agarose gel. The fractions containing fragments of suitable size are pooled and concentrated for further processing. There may be regarded as suitable within the scope of this invention especially fragments of an order of magnitude of from 30 kb to 45 kb, preferably from 40 kb to 45 kb.

In parallel with the fragmentation described above, or later, for example for the subsequent ligase reaction, a suitable cosmid vector, such as pHC79 (Hohn & Collins, Gene (1980), **11**, 291) or pWE15® (Stratagene) is completely digested with a suitable restriction enzyme, such as BamHI.

The ligation of the cosmid DNA with the *Streptomyces* or *Actinomyces* fragments fractionated according to size can be carried out using a T_4-DNA ligase. After an adequate incubation period, the ligation batch so obtainable is packaged into λ-phages by generally known methods.

The resulting phage particles are then used to infect a suitable host strain. Preference is given to a recA⁻ *E. coli* strain, such as. *E. coli* HB101 or X-1 Blue® (Stratagene). The selection of transfected clones and the isolation of the plasmid DNA can be carried out using generally known methods.

Screening of the gene bank for DNA fragments that play a role in staurosporin biosynthesis is carried out using a specific hybridisation probe which is assumed (for example on the basis of complementation tests or gene disruption) to contain DNA regions of the staurosporin gene cluster.

Differential screening of the resulting transformed colonies can be used to detect suitable colonies and to isolate their plasmid DNA (Maniatis *et al.*, 1982; pp. 368-369). The isolated plasmid DNA is then cleaved with a suitable restriction enzyme and analysed by means of agarose gel electrophoresis for the size of the inserted fragments, the previously selected plasmid PSLO18/10 being used, for example, as reference standard.

A plasmid containing an additional fragment of the desired size can then be isolated from the gel in the manner described hereinbefore. Confirmation that the additional fragment is identical to the desired fragment of the previously selected cosmid can then be obtained by means of Southern transfer and hybridisation.

Analysis of the function of the DNA fragments thus isolated can be carried out within the context of a gene disruption experiment, as described hereinbefore.

The invention relates also to the use of the DNA fragments, hybrid vectors, expression cassettes or transformed host organisms according to the invention in the preparation of indole-carbazole alkaloids and especially of staurosporin and its precursors or derivatives.

Derivatives of staurosporin are customarily understood as being those having modified substitution patterns which either serve as the starting point for further modifications or can themselves be used as active ingredients or prodrugs.

The DNA fragments, hybrid vectors or expression cassettes according to the invention can be used both in the preparation of indole-carbazole alkaloids, and especially staurosporin, in host organisms not previously capable of producing indole-carbazole alkaloids and to improve the yield in organisms already producing indole-carbazole alkaloids. For that purpose, for example, a plurality of copies of relevant DNA fragments can be inserted into the host organisms, or the regulatory mechanisms of indole-carbazole alkaloid biosynthesis, and especially of staurosporin biosynthesis, can be analysed and modified in order to improve production. It is also possible, by combining DNA fragments from indole-carbazole alkaloid gene clusters with other DNA fragments, for example, to replace specific enzymes, in order to produce derivatives of those alkaloids.

A further possible use of the DNA fragments according to the invention consists in inactivating enzymes that are involved in indole-carbazole alkaloid biosynthesis or in using the DNA fragments according to the invention in the synthesis of oligonucleotides which are then used in the context of PCR amplification to detect homologous sequences.

Figures

Fig. 1 10 kb DNA region containing a number of important restriction cleavage sites
Fig. 1 35 kb DNA region containing a number of important restriction cleavage sites

Examples

All liquid cultures of *S. longisporoflavus* are carried out in Erlenmeyer flasks at 28°C or 30°C on a shaker at 250 rpm. General molecular genetic techniques, such as agarose gel electrophoresis, restriction digestion, DNA purification by ethanol precipitation, and DNA isolation from agarose, are carried out as described in Maniatis *et al.*, Molecular Cloning: A laboratory manual, 1st Edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor NY

(1982) or in *Sambrook et al.*, Molecular Cloning: A laboratory manual, 2nd Edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor NY, 1989.

Nutrients used:

LB	Maniatis *et al.*, Molecular Cloning: A laboratory manual, 1st Edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor NY (1982)
TSB medium	Hopwood *et al.* (Genetic manipulation of *Streptomyces*, a laboratory manual. The John Innes Foundation, Norwich (1985))
minimal agar (MM)	Hopwood *et al.* (Genetic manipulation of *Streptomyces*, a laboratory manual. The John Innes Foundation, Norwich (1985))
R2YE agar plate	Hopwood *et al.* (Genetic manipulation of *Streptomyces*, a laboratory manual. The John Innes Foundation, Norwich (1985))
DST (=SNA) soft agar	Hopwood *et al.* (Genetic manipulation of *Streptomyces*, a laboratory manual. The John Innes Foundation, Norwich (1985))
NL148 (=NL148G without glycine)	Schupp *et al.* FEMS Microbiology Lett. (1986), **36**, 159-162
NL19Q	Schupp *et al.* FEMS Microbiology Lett. (1987), **42**, 135-139
SCR12mod	20 g/l full-fat soya flour 20 g/l saccharose 12 g/l HEPES 0.1 g/l SAG 471 antifoam adjust pH to 7.5 with NaOH before sterilisation (autoclaving)
SET	75 mM NaCl, 25 mM EDTA, 20 mM Tris, pH 7.5

Example 1: Obtaining high-molecular-weight genomic DNA fragments from *S. longisporoflavus*

In order to obtain high-molecular-weight genomic DNA from *S. longisporoflavus*, cells of the strain *S. longisporoflavus* R19 DSM 10189 are cultured for 24 hours at 28°C in SCR12mod medium. 5 ml of the culture are then transferred to 100 ml of NL148 medium (+ 2.5 g/l glycine) in a 500 ml Erlenmeyer flask and the culture is incubated for 48 hours at 28°C. The cells are separated from the medium by centrifuging at 3000 g for 10 min. and are resuspended in 5 ml of SET (75 mM NaCl, 25 mM EDTA, 20 mM Tris, pH 7.5). The

extraction of high-molecular-weight chromosomal DNA is effected in accordance with the method of A. Pospiech and J. Neumann (Trends in Genetics (1995), **11**, 217-218).

The high-molecular-weight genomic DNA of *S. longisporoflavus* thus isolated is partially digested in portions of approximately 5 µg of DNA using the restriction enzyme Sau3A (Boehringer, Mannheim), forming DNA fragments the majority of which are from 5 to 40 kb in size. The requisite amount of enzyme, in a range of from 0.002 - 0.02 units/µg of DNA, is determined empirically by analysis of the digestion (37°C, 30 minutes) using agarose gel electrophoresis. The enzyme reaction is stopped by incubation for 15 minutes at 65°C, followed by phenol/chloroform extraction and ethanol precipitation.

The DNA thus pretreated is separated according to fragment size by centrifuging (83000 g, 20°C) for 18 hours over a 10% to 40% saccharose density gradient. The gradient is fractionated in aliquots of 0.5 ml and dialysed. 10 µl samples are analysed on a 0.3% agarose gel using a DNA size standard. Fractions containing chromosomal DNA of the desired size are collected, precipitated with ethanol and concentrated.

Example 2: Cloning of random DNA fragments of *S. longisporoflavus* R19 (DSM 10189) into plasmid vector pIJ486

For cloning *S. longisporoflavus* DNA fragments, the generally known *Streptomyces* plasmid vector pIJ486, which has a wide range of hosts and is present in a large number of copies per cell (Ward *et al.*, Mol. Gen. Genet. (1986), **203**, 468-478), is used. The vector is first transformed into *S. longisporoflavus* R19 using the general transformation conditions for *Streptomyces* described in Hopwood *et al.* (Genetic manipulation of *Streptomyces*, a laboratory manual. The John Innes Foundation, Norwich (1985) pages 110-111). For further work with *S. longisporoflavus*, the plasmid pIJ486 is isolated from *S. longisporoflavus* using a CsCl preparation. For that purpose, cells of *S. longisporoflavus* containing pIJ486 are cultured for 48 hours at 28°C in NL19Q medium. Then 10 x 2.5 ml of culture are used to inoculate 200 ml Erlenmeyer flasks with 50 ml of nutrient solution NI148 each, and incubated for 48 hours at 28°C. pIJ486 plasmid DNA is then isolated from the 500 ml of culture solution; Hopwood *et al.* (pages 82-84).

In order to clone S. longisporoflavus DNA fragments, the vector pIJ486 is cleaved completely with the restriction enzyme BamHI, precipitated with ethanol and then treated with alkaline phosphatase (Boehringer, Mannheim) in accordance with the manufacturer's instructions, in order to prevent self-ligation of the plasmid in the subsequent ligation

reactions. The vector thus treated is ligated with partially Sau3A-digested chromosomal DNA of S. longisporoflavus (fraction after sucrose gradient with DNA fragments of 5-20 kb, see above). The ligation is effected with T4-DNA ligase (Boehringer, Mannheim) in accordance with the manufacturer's instructions and with approximately equimolar amounts of the two DNA starting materials and a final concentration of total DNA of approximately 600 mg/ml in a ligation volume of 10 ml. 1 ml of the ligation batch is used to transform the *S. longisporoflavus* mutant M14, which is blocked in the final step of staurosporin biosynthesis and produces the staurosporin analogue 3'-demethyl-3'hydroxystaurosporin (Hoehn *et al.*, J. Antibiotics (1995), **48**, 300-305), using the general transformation conditions for Streptomyces described in Hopwood *et al.* (pages 110-111). The transformation batch is then plated out onto R2YE agar (Hopwood *et al.*, page 236). In order to select the colonies containing the plasmid, after 20 hours 30 µg/ml of thiostreptone (final concentration) are poured over the plates. For the plasmid preparation, 24 thiostreptone-resistant colonies are each transferred into 25 ml of TSB medium containing 30 µg/ml of thiostreptone (50 ml Erlenmeyer flasks, each containing 10-20 sterile quartz splinters in order to produce short mycelium fragments) and incubated for 48 hours at 28°C. The plasmids are then isolated from those cultures using a slight modification of the method of Birnboim and Doly (Nucl. Acids Res. (1979), **7**, 1513-1523). The method is modified as follows: lysozyme digestion for 60 minutes at 30°C in the following solution: 2 mg/ml of lysozyme, 10 mM EDTA, 25 mM tris pH 8.0, 10% glucose). Analysis of the plasmids shows that approximately 60% of the transformed colonies contain a 5-20 kb DNA fragment, integrated in the plasmid.

Example 3: Identification and cloning of a *S. longisporoflavus* DNA fragment that complements the blocked mutant M14 clone for normal staurosporin production

12 300 transformed colonies of the mutant *S. longisporoflavus* M14 are obtained in several series from the ligation batch described above and analysed for complementation of the blocked staurosporin biosynthesis step. From the investigations carried out above, it can be inferred that 60%, or approximately 7380, of the clones investigated contain plasmid pIJ486, together with an additional DNA fragment of *S. longisporoflavus*. After incubating the R2YE plates at 28°C for 6 days, the 12 300 colonies are screened (pretested) as follows in a biological test for staurosporin production:

Biological test: In order to transfer (replica plate) all 12 300 colonies to a different agar, sterile Whatman W541 filter paper is placed on each R2YE agar plate, the plate is incubated overnight at 28°C and the filter is then removed in a sterile manner and placed

carefully on plates containing MM minimal agar (Hopwood *et al.*, page 233). After incubation of the MM plates for 24 hours at 28°C, the filter paper is removed and the plates are incubated for a further 24 hours. Using that procedure, the colonies are transferred 1 to 1 from the original R2YE agar to the MM agar, the original R2YE plates serving at the same time as original plates for the further processing of colonies that exhibit positive results in the biological test. 6 ml of DST soft agar (48°C) containing approximately 10^7 cells of *Saccharomyces cerevisiae* ATCC 9763 are poured over each of the MM plates which contain small but visibly replicated colonies. Those plates are incubated overnight at 30°C and then investigated for inhibition zones (in lawns where the *Saccharomyces cerevisiae* test organism has grown) produced by the *S. longisporoflavus* colonies. Under those test conditions, colonies of *S. longisporoflavus* R19 produce an inhibition zone 2-4 mm in diameter as a result of their staurosporin production, whereas colonies of the blocked mutant M14 do not normally produce an inhibition zone.

In one colony, a significant inhibition zone can be detected using this biological test. That clone is isolated from the original R2YE plate and, in order to isolate the plasmid, transferred into 25 ml of TSB medium containing 30 µg/ml of thiostreptone (50 ml Erlenmeyer flasks, each containing 10-20 sterile quartz splinters in order to produce short mycelium fragments) and incubated for 48 hours at 28°C. The plasmid DNA is then isolated from the culture using a slight modification of the method of Birnboim and Doly (Nucl. Acids Res. (1979), **7**, 1513-1523) (see above) and analysed. The clone contains small amounts of recombinant plasmid DNA, together with an additional *S. longisporoflavus* DNA fragment of approximately 20 kb. This plasmid preparation is given the number pSLO18/10.

In order to monitor the complementation of the blocking of the mutant *S. longisporoflavus* M14 by the plasmid DNA pSLO18/10, the latter is again transformed into that mutant. It is now found that 3 out of 10 transformed M14 colonies are complemented by the plasmid DNA for approximately normal staurosporin production. The plasmids of those 3 clones are identical (number pSLO18/10/2) and contain an inserted DNA fragment of approximately 20 kb in which an internal 2.1 kb BglII fragment is detectable.

Example 4: Analysis of the cloned 2.1 kb BglII DNA fragments

The first step is to determine whether the identified and cloned 2.1 kb BglII fragment of *S. longisporoflavus* is sufficient alone to complement the *S. longisporoflavus* mutant M14. For that purpose, the DNA fragment is isolated from the plasmid pSLO18/10/2, subcloned into the vector pIJ486 and transformed into the *S. longisporoflavus* M14 mutant. Analysis of the

clones thus obtained reveals that all of the clones that contain the 2.1 kb BglII DNA fragment are complemented for normal staurosporin production (approximately equivalent to the parent strain R19 of the mutant M14). Cultures in a liquid medium give HPLC values of 100 - 200 mg/l staurosporin, while the value for the mutant M14 is 0-5 mg/l. The plasmid containing the 2.1 kb BglII fragment of *S. longisporoflavus* is given the number pSLO24/3.

In order to demonstrate that the 2.1 kb BglII fragment is a chromosomal *S. longisporoflavus* DNA fragment, the fragment is radioactively labelled with $CTP^{32}P$ (see below) and analysed as a probe in a Southern Blot with BglII-digested chromosomal DNA of *S. longisporoflavus* R19. The experiment confirms that the cloned 2.1 kb BglII fragment is an authentic chromosomal fragment of *S. longisporoflavus* R19.

Example 5: DNA sequence determination of the 2.1 kb BglII fragment

For sequencing, the 2.1 kb BglII fragment is first isolated from the plasmid pSLO24/3 (Maniatis *et al.*, 1982) and subcloned into the BamHI cleavage site of the vector pUC18 which is suitable for DNA sequencing (pSL26/1 = number of the new plasmid). In addition, a 1.1 kb SalI subfragment, which is located internally in the 2.1 kb BglII fragment, is cloned into vector pUC18 in both orientations (pSLO32/13, pSLO32/19). The DNA of the three plasmids pSLO26/1, pSLO32/13, pSLO32/19 is sequenced using the dideoxy nucleotide chain-termination method of Sanger, with stain-labelled primers, and the Applied Biosystems automatic sequencer (Model 373A) in accordance with the manufacturers' instructions. Universal pUC18 primers and new oligonucleotide primers, constructed in accordance with newly obtained sequences in the two BglII and SalI fragments, are used for the double-stranded sequencing. The resulting DNA sequences from individual runs are assembled using Applied Biosystems software. In that manner, both DNA strands of the the 2.1 kb BglII fragment of *S. longisporoflavus* R19 can be fully sequenced. The DNA sequence of the 2.1 kb BglII fragment, which is 2122 base pairs in length, is set out in SEQ ID NO 1.

Example 6: Analysis of 2 regions (genes) coding for proteins on the 2.1 kb BglII fragment

The nucleotide sequence of the 2.1 kb BglII fragment is analysed using the computer program Codonpreference (Genetics Computer Group 1994). The analysis shows that two distinct open reading frames (ORF), each coding for one protein, are present. The codons used in the two ORFs are typical for *Streptomyces* genes, from which it may be deduced that there are two genes on the 2.1 kb BglII fragment of *S. longisporoflavus*. A comparison

of the two genes of *S. longisporoflavus* and of the proteins derived therefrom with DNA/-protein sequences from the GenBank/EMBL data bank yields the following results:

Gene 1 (ORF of base pair 845 - 1684; SEQ ID NO 2): codes for a protein containing 280 amino acids. The protein is significantly similar to known S-adenosyl methionine-dependent methyl transferases, especially to those of *Streptomyces* and *Actinomyces*, which are involved in the transfer of methyl groups in secondary metabolite biosyntheses. In particular, the protein derived from gene 1 has the three typical sequence motifs that are characteristic of such methyl transferases. A comparison of the motif 1 sequences is given here as an example:

Microorganism	Gene	Sequence	Product
S. longisporoflavus		**VLDLGCGVG**	staurosporin O-MT
S. erythraea	eryG	**VLDVGFGLG**	erythromycin O-MT
S. peuceticus	dnrK	**VLDVGGGKG**	carminomycin O-MT
S. mycarofaciens	mdmC	**VLEIGTGTG**	midamycin O-MT
S. glaucescens	tcmO	**FVDLGGARG**	tetracenomycin O-MT
Consensus O-MT general		**VLDIGGGTG**	

As demonstrated above, the 2.1 kb BglII fragment of *S. longisporoflavus* is capable of complementing the mutant M14 which is blocked in precisely such a methyl transferase step in the biosynthesis of staurosporin. That finding, together with the sequence analysis, which showed significant homology between the gene product of gene 1 of the BglII fragment and methyl transferases, leads to the definite conclusion that gene 1 codes for a methyl transferase that is responsible for the O-methylation step from 3'-demethoxy-3'-hydroxystaurosporin to staurosporin in the biosynthesis of staurosporin.

Gene 2 (ORF of base pair 148 - 768; SEQ ID NO: 3): codes for a protein containing 207 amino acids. The protein is significantly similar to the dTDP-4-keto-6-deoxyglucose 3,5-epimerase of *Streptomyces glaucescens*, that is to say there is 48.6% amino acid identity over a region of 148 amino acids. The dTDP-4-keto-6-deoxyglucose 3,5-epimerase of *Streptomyces* is involved in the synthesis of the deoxy sugar moiety of metabolites, such as streptomycin. Since staurosporin also has a deoxy sugar moiety in the molecule, it may be concluded that gene 2 of the 2.1 kb BglII fragment is involved in the synthesis of that moiety of the staurosporin molecule.

The above assumption regarding gene 2 made as a result of the sequence comparison can be confirmed by the fact that the *S. longisporoflavus* mutant M13 (Hoehn *et al.*, J. Antibiotics (1995), **48**, 300-305), which is blocked in a synthesis step of the deoxy sugar moiety of staurosporin, can be complemented for normal staurosporin production by the 2.1 kb BglII fragment. Gene 2 of the 2.1 kb fragment of *S. longisporoflavus* is thus involved in a biosynthesis step in the deoxy sugar moiety of staurosporin.

Example 7: Construction of a cosmid gene bank of *S. longisporoflavus* R19

The commercially available plasmid pWE15 (Stratagene, La Jolla, CA, USA) is used as the cosmid vector. pWE15 is cleaved completely using the enzyme BamHI (Maniatis *et al.* 1989) and precipitated with ethanol. The cosmid DNA is ligated with the corresponding size-fractionated *S. longisporoflavus* Sau3A DNA fragments (see above) with the aid of a T4-DNA ligase. During the ligation, approximately 3 µg each of the two DNA starting materials are used in a reaction volume of 20 µl, and the ligation is carried out for 15 hours at 12°C.

Using the *in vitro* packaging kit commercially available from Stratagene (La Jolla, CA, USA), 4 µl of the above ligation batch are packaged in lambda phages (in accordance with the manufacturer's instructions). The resulting phages are introduced into the *E. coli* strain X-1BlueMR® (Stratagene) by means of infection. Titration of the phage material yields approximately 20 000 phage particles per ml and an analysis of 12 cosmid clones shows that all the clones contain 30 - 40 kb plasmid DNA inserts.

Example 8: Preparation of a radioactive probe of the 2.1 kb BglII fragment of *S. longisporoflavus*

The plasmid pSL26/1, which contains the 2.1 kb BglII fragment in the *E. coli* vector pUC18, is used as the starting material for the preparation of the DNA probe. The 2.1 kb insert fragment is separated off by means of EcoRI + HindIII digestion and then separated using agarose gel. Approximately 1 µg of the isolated 2.1 kb DNA fragment is radioactively labelled with ^{32}P-d-CTP by means of the nick-translation system from GIBCO/BRL (Basle) in accordance with the manufacturer's instructions.

Example 9: Isolation of four cosmid clones with chromosomal *S. longisporoflavus* DNA fragments containing the 2.1 kb BglII fragment

By infection of *E. coli* X-1 Blue MR® (Stratagene) with an aliquot of the *in vitro*-packaged lambda phages (see above), over 4000 clones are obtained on a plurality of LB + ampicillin + neomycin plates (50 µg/ml of each). The clones are tested by colony hybridisation on nitrocellulose filters (Schleicher + Schuell). The ^{32}P-d-CTP radioactively labelled 2.1 kb *S. longisporoflavus* fragment prepared above is used as DNA probe.

6 cosmid clones are found that exhibit a significant signal with the DNA probe. The plasmid DNA of those cosmids is isolated (Maniatis *et al.* 1989), digested with BglII and analysed in an agarose gel. The analysis shows that all 4 recombinant plasmids contain inserted chromosomal *S. longisporoflavus* DNA approximately 35 kb in size and all 6 contain the 2.1 kb BglII fragment.

Example 10: Characterisation of the chromosomal *S. longisporoflavus* DNA region adjacent to the cloned BglII fragment

In order to characterise the chromosomal *S. longisporoflavus* DNA region adjacent to the cloned BglII fragment, a restriction analysis of the plasmid DNA of one of the 6 cosmid clones is carried out. The selected plasmid of the cosmid clone has the number pNE29 (DSM 10188).

In order to identify the fragments that overlap with the BglII fragment, the plasmid pNE29 is digested with enzymes EcoRI, PvuI, PvuII and BclI and tested in a Southern Blot (Maniatis *et al.* 1989) with the 2.1 kb fragment as probe. The result of the analysis is that in each case 2 or 1 DNA fragment(s) of the following size overlap(s) with the 2.1 kb fragment: EcoRI: > 20 kb, PvuII: 3.5 kb and 6.5 kb; PvuI: 3.4 kb and 2.1 kb; BclI: 3.6 kb. An approximately 10 kb DNA region of the chromosome of *S. longisporoflavus* can thus be characterised (Figure 1).

By means of a further restriction analysis of the plasmid pNE29, a rough restriction map of that region of the *S. longisporoflavus* chromosome can be prepared which allows the approximately 35 kb DNA region to be characterised. The restriction map is shown in Figure 2.

Example 11: DNA sequence determination of the 6 kb PvuII-BglII fragment immediately preceding the sequenced 2.1 kb BglII fragment (see Figure 1)

The 6 kb PvuII-BglII fragment immediately preceding the sequenced 2.1 kb BglII fragment (on the left in Figure 1) is sequenced using the 6.5 kb PvuII fragment from the approximately 10 kb region of the *S. longisporoflavus* chromosome characterised in Example 10. For that purpose, the 6.5 kb PvuII fragment is isolated from cosmid pNE29 or cosmid pNE31 (one of the 4 cosmids from Example 9), which is identical in that region (Maniatis *et al.*, 1982), and subcloned into the SmaI cleavage site of the vector pBluescript II SK (Stratagene) suitable for DNA sequencing (pNE37 = number of the new plasmid). In addition, SmaI subfragments located internally in the 6.5 kb PvuII fragment are cloned into the SmaI cleavage site of the vector pBluescript II SK. The DNA sequencing is effected with the plasmids using the dideoxy nucleotide chain-termination method of Sanger, as described in Example 5. Universal pBluescript primers and new oligonucleotide primers, constructed in accordance with newly obtained DNA sequences, are used for the double-stranded DNA sequencing. The resulting DNA sequences are joined together and analysed using software from Applied Biosystems and the Genetics Computer Group (1994). In that manner the complete DNA sequence of the 6 kb PvuII-BglII fragment of *S. longisporoflavus* can be determined. That DNA sequence is set out in SEQ ID NO 4. The resulting sequence of the 0.5 kb BglII-PvuII region of the 6.5 kb PvuII fragment shows that the two DNA sequences SEQ ID NO 1 and SEQ ID NO 4 of *S. longisporoflavus* are connected to one another directly *via* the BglII cleavage site.

Example 12 : Analysis of 5 regions (genes) coding for proteins on the 6.5 kb PvuII fragment of *S. longisporoflavus* (see Fig. 1)

The nucleotide sequence of the 6.5 kb PvuII fragment is analysed using the computer program Codonpreference (Genetics Computer Group 1994). The analysis shows that 5 distinct open reading frames (ORF) that code for proteins are present. The codons used in the ORFs are typical for *Streptomyces* genes, from which it can be deduced that there are 5 genes on the 6.5 kb PvuII fragment of *S. longisporoflavus*. A comparison of the 5 genes of *S. longisporoflavus* and the proteins derived therefrom with DNA/protein sequences from the gene/EMBL data bank yields the following results:

Gene 1 (ORF of base pair 378 - 1655 of SEQ ID NO 4) codes for a protein containing 425 amino acids.

Gene 2 (ORF of base pair 1747 - 2553 of SEQ ID NO 4) codes for a protein containing 268 amino acids. The protein is significantly similar to known S-adenosyl methionine-dependent methyl transferases, especially to those of *Streptomyces* and *Actinomyces*, which are involved in the transfer of methyl groups to secondary metabolites. On the basis of that similarity it can be concluded that the methyl transferase is involved in the N-methylation step of the sugar in staurosporin biosynthesis.

Gene 3 (ORF of base pair 2593 - 4011 of SEQ ID NO 4) codes for a protein containing 472 amino acids.

Gene 4 (ORF of base pair 4013 - 4999 of SEQ ID NO 4) codes for a protein containing 328 amino acids.

Gene 5 (ORF of base pair 5071 - 6171 of SEQ ID NO 4) codes for a protein containing 366 amino acids. That protein is significantly similar to amino transferase enzymes, such as the DnrJ protein of *Streptomyces peuceticus*. Those enzymes, which are involved in the biosynthesis of antibiotics, are ascribed the function of adding an amino group in the biosynthesis of the deoxyamino sugar moiety of the antibiotic. On the basis of that similarity, it can be concluded that gene 5 is involved in the synthesis of the deoxyamino sugar in the biosynthesis of staurosporin.

Example 13 : DNA sequence determination of the 1.8 kb BglII - PvuII region immediately following the sequenced 2.1 kb BglII fragment (corresponds to the right-hand BglII - PvuII end fragment in Figure 1)

The approximately 1.8 kb BglII - PvuII region to the right of the sequenced 2.1 kb BglII fragment (Figure 1) is sequenced using the 3.5 kb PvuII fragment from the approximately 10 kb region of the *S. longisporoflavus* chromosome characterised in Example 10. For that purpose, the 3.5 kb PvuII fragment is isolated from cosmid pNE29 or cosmid pNE31 (one of the 4 cosmids from Example 9), which is identical in that region (Maniatis *et al.*, 1982), and subcloned into the SmaI cleavage site of the vector pBluescript II SK (Stratagene) which is suitable for DNA sequencing (pNE36 = number of the new plasmid). In addition, SmaI subfragments located internally in the 3.5 kb PvuII fragment are cloned into the SmaI cleavage site of the vector pBluescript II SK. The DNA sequencing is carried out with the plasmids using the dideoxy nucleotide chain-termination method of Sanger, as described in Example 5. Universal pBluescript primers and new oligonucleotide primers, constructed in

accordance with newly obtained DNA sequences, are used for the double-stranded DNA sequencing. The resulting DNA sequences are joined together and analysed using software from Applied Biosystems and the Genetics Computer Group (1994). In that manner the complete DNA sequence of the 1.8 kb BglII - PvuII region of *S. longisporoflavus* can be determined. The overlaps between the resulting sequences of the whole 3.5 kb PvuII fragment used for the sequencing and SEQ ID NO 1 (2.1 kb BglII fragment) show that the 2.1 kb BglII and 1.8 kb BglII-PvuII DNA regions of *S. longisporoflavus* shown in Figure 1 are connected not directly, but *via* a BglII fragment having only 69 base pairs. The entire DNA sequence from immediately adjacent to the right-hand side of the 2.1 kb BglII fragment to the next PvuII cleavage site (right-hand end in Figure 1) is set out in SEQ ID NO 5. Taken together, the DNA sequences SEQ ID NO 4, SEQ ID NO 1 and SEQ ID NO 5 thus represent the DNA sequence of the region of *S. longisporoflavus* shown in Figure 1.

Deposited microorganisms

The following microorganisms and plasmids have been deposited with the "Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSM)" (German Collection of microorganisms and cell cultures), Mascheroder Weg 1b, D-38124 Brunswick, in accordance with the requirements of the Budapest Convention:

Microorganism/plasmid	**Date of deposition**	**Deposit number**
Streptomyces longisporoflavus	23.08.95	DSM 10189
E. coli/pNE29	23.08.95	DSM 10188

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13*bis*)

A. The indications made below relate to the microorganism referred to in the description on page 29, line 1-8.

B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet ☐

Name of depositary institution: Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSM)

Address of depositary institution *(including postal code and country)*: Mascheroder Weg 1B, D-38124 Braunschweig, Germany

Date of deposit	Accession Number
23 August 1995 (23.08.95)	DSM 10188

C. ADDITIONAL INDICATIONS *(leave blank if not applicable)* This information is continued on an additional sheet ☐

We request the Expert Solution where available

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE *(if the indications are not for all designated States)*

E. SEPARATE FURNISHING OF INDICATIONS *(leave blank if not applicable)*

The indications listed below will be submitted to the International Bureau later *(specify the general nature of the indications e.g., "Accession Number of Deposit")*

For receiving Office use only	For International Bureau use only
☒ This sheet was received with the international application	☐ This sheet was received by the International Bureau on:
Authorized officer: R.L.R. Pether	Authorized officer

Form PCT/RO/134 (July 1992)

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:

(A) NAME: CIBA-GEIGY AG

(B) STREET: Klybeckstr. 141

(C) CITY: Basle

(E) COUNTRY: Switzerland

(F) POSTAL CODE (ZIP): 4002

(G) TELEPHONE: +41 61 69 11 11

(H) TELEFAX: + 41 61 696 79 76

(I) TELEX: 962 991

(ii) TITLE OF INVENTION: Staurosporin biosynthesis gene clusters

(iii) NUMBER OF SEQUENCES: 5

(iv) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk

(B) COMPUTER: IBM PC compatible

(C) OPERATING SYSTEM: PC-DOS/MS-DOS

(D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2122 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: misc_RNA

(B) LOCATION:1..2122

(D) OTHER INFORMATION:/product= "2.1 kb region"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

GGATCTTCTC GCTGCCGATG TACCCCTCGC TCGCCCCCGA CCTCCAGGAC AAGGTCATCC 60

ACGCCGTACG CGAGGTGCTC GCCACTCTGT GACTGTCCGT CAACTCTCTT ATCGCATCGC 120

GTCGTTCACC GAGTCACTGG AGCAGAAGTG AAAGCACGCC CGCTCACCGT CGAGGGAGCC 180

GTCGAGTTCA CCCCCCGCGT CTTCCCCGAC GACAGGGGCA AGTTCGTCTC GCCGTACCAG 240

GAAGCGACGT TCACCGAGGC CCACGGCACC CCGCTCTTCC CCGTGGCGCA GACCAACCAC 300

AGCGTGTCCC GGCGAGGTGT CGTACGCGGC GTCCACTACA CGGCGACGCC CCCGGGCACC 360

GCCAAGTACG TCTACTGCGC CCGAGGCCGC GCCCTGGACA TCGTCGTCGA CATCCGCGTC 420

GGCTCGCCCA CCTTCGGCCG CTGGGACGCG GTGCTGATGG ACCAGCTGGA TCACCGGGCC 480

AGCTATTTTC CCGTCGGGGT CGGCCATGCC TTCGTGGCCC TGGAGGACGA CACCGACATG 540

TCGTACATGC TCTCCGGGCG CTATGTCGCC GAGCACGAAC TCTCCCTGTC CGCCCTGGAC 600

CCGGACCTCG GGCTGCCGAT CCCCACGGAC CTCGAACCGA TCCTCTCCGA ACGCGACCGC 660

GCGGCCGTCA CCCTCGCCGA GGCCCAGGAG AAGGGCCTGC TGCCGGACTA CGCCCGCTGC 720

CAGGAGATCG AGCGGGGACT CGTCCCCCGC GCGAGGCCGG CGGCGTAGCC CCGCACCGAC 780

GAGGCATTTC ACTCCCCTTC TCACTCCCTT TCTCACTGTC GATCGATCCG AAAGGCCGTT 840

CCCATGACCG ACTCCACCCA GACCCTGCCC GTGCCGGAAG CCGTCGGTGA GCTGTACGAC 900

CGGCTGACGC TGAGCGCGAT GAACGACGGC TCGTTCAACC CCAATGTGCA CATCGGCTAT 960

TGGGACACCC CGGGCTCCGA GGCCACCATC GAGGAGGCGA TGGACCGGCT CACCGATGTG 1020

TTCATCGAAC GGCTGAACGC GTACGCCACC TCCCACGTCC TCGACCTCGG CTGCGGGGTG 1080

GGCGGGCCCG GCCTCAGGGT CGTGGCGCGC ACCGGGGCAC GGGTCACCGG CATCAGCATC 1140

AGCGAGGAGC AGATCAGGAC CGCCAACCGG CTGGCCGCCG AGGCCGGGGT CGCCGACCGT 1200

GCCGTGTTCC AGCATGGCGA CGCGATGAAA CTGCCCTTCG CCGACGCCTC GTTCGACGCC 1260

GTGATGGCGC TGGAGTCGAT CTGCCACATG CCCGACCGGC AGCAGGTGTT CACCGAGGTG 1320

TGCCGGGTGC TGCGCCCCGG GGGCCGGATC GTCCTCACCG ACATCTTCGA GCGCCACCCG 1380

CGCAAGGCGG TACGACACCC CGGCATCGAC AAGTTCTGCC GCGACCTGAT GTCGACCACG 1440

GCGGACATCG ACGACTACGT GGCGCTGCTG CACCGCTCCG GGCTGCGGCT GCGCGAGATC 1500

GTCGACGTCA CCGAGCAGAC CACGCTGCGC CTCGCCGACG AGATCGGCAG GCTCGCGGCC 1560

GTCGAGGAGC GCCCCGTGGC CATGGACGAG GGCAACTTCG CCTTCGGCGA CGACTCCTTC 1620

AAGCCGTCCG ACCTGGCGGG CGTCGACGAC TTCGGCTGCC TCCTGGTCAC CGCCGAGCGC 1680

CCCTGACCCG CTGAAACGCC GGGAGGTCAG GCGCACCTGC CCTCCCGGCG CCCGTCCCCC 1740

GGGTCGCGAG CGCATTGCAT CCCCCGTGCC GCGAGCCCAC GCATTCCCCG GGCCACGAGC 1800

CCACGCGTCC GCGACACGGA CCCACAAGGA GAGGCAAGAA CGAGATGACG CATTCCGGTG 1860

AGCGGACCGA TGTGCTGATC GTGGGCGGCG GCCCGGTCGG GATGGCGCTG GCGCTGGATC 1920

TGAGGTACCG GGGCATCGAC TGTCTGGTCG TCGACGCCGG TGACGGCACG GTCCGGCACC 1980

CCAAGGTCAG CACCATCGGT CCCCGCTCGA TGGAACTCTT CCGCCGCTGG GGCGCCGCGG 2040

ACGCGATCCG GAACGCCGGC TGGCCCGCCG ACCATCCCCT GGACATCGCC TGGGTGACCA 2100

AGGTCGGCGG CCACGAAGAT CC 2122

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 280 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:
(A) NAME/KEY: Protein
(B) LOCATION:1..280
(D) OTHER INFORMATION:/note=
"methyl transferase-like protein"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Met Thr Asp Ser Thr Gln Thr Leu Pro Val Pro Glu Ala Val Gly Glu
1 5 10 15

Leu Tyr Asp Arg Leu Thr Leu Ser Ala Met Asn Asp Gly Ser Phe Asn
20 25 30

Pro Asn Val His Ile Gly Tyr Trp Asp Thr Pro Gly Ser Glu Ala Thr
35 40 45

Ile Glu Glu Ala Met Asp Arg Leu Thr Asp Val Phe Ile Glu Arg Leu

50 55 60

Asn Ala Tyr Ala Thr Ser His Val Leu Asp Leu Gly Cys Gly Val Gly
65 70 75 80

Gly Pro Gly Leu Arg Val Val Ala Arg Thr Gly Ala Arg Val Thr Gly
85 90 95

Ile Ser Ile Ser Glu Glu Gln Ile Arg Thr Ala Asn Arg Leu Ala Ala
100 105 110

Glu Ala Gly Val Ala Asp Arg Ala Val Phe Gln His Gly Asp Ala Met
115 120 125

Lys Leu Pro Phe Ala Asp Ala Ser Phe Asp Ala Val Met Ala Leu Glu
130 135 140

Ser Ile Cys His Met Pro Asp Arg Gln Gln Val Phe Thr Glu Val Cys
145 150 155 160

Arg Val Leu Arg Pro Gly Gly Arg Ile Val Leu Thr Asp Ile Phe Glu
165 170 175

Arg His Pro Arg Lys Ala Val Arg His Pro Gly Ile Asp Lys Phe Cys
180 185 190

Arg Asp Leu Met Ser Thr Thr Ala Asp Ile Asp Asp Tyr Val Ala Leu
195 200 205

Leu His Arg Ser Gly Leu Arg Leu Arg Glu Ile Val Asp Val Thr Glu
210 215 220

Gln Thr Thr Leu Arg Leu Ala Asp Glu Ile Gly Arg Leu Ala Ala Val
225 230 235 240

Glu Glu Arg Pro Val Ala Met Asp Glu Gly Asn Phe Ala Phe Gly Asp
245 250 255

Asp Ser Phe Lys Pro Ser Asp Leu Ala Gly Val Asp Asp Phe Gly Cys
260 265 270

Leu Leu Val Thr Ala Glu Arg Pro
275 280

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 206 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:
(A) NAME/KEY: Protein
(B) LOCATION:1..206
(D) OTHER INFORMATION:/note= "NDP-4-keto-6-deoxyhexose 3,5-epimerase-like protein"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

Val Lys Ala Arg Pro Leu Thr Val Glu Gly Ala Val Glu Phe Thr Pro
1 5 10 15

Arg Val Phe Pro Asp Asp Arg Gly Lys Phe Val Ser Pro Tyr Gln Glu
20 25 30

Ala Thr Phe Thr Glu Ala His Gly Thr Pro Leu Phe Pro Val Ala Gln
35 40 45

Thr Asn His Ser Val Ser Arg Arg Gly Val Val Arg Gly Val His Tyr
50 55 60

Thr Ala Thr Pro Pro Gly Thr Ala Lys Tyr Val Tyr Cys Ala Arg Gly
65 70 75 80

Arg Ala Leu Asp Ile Val Val Asp Ile Arg Val Gly Ser Pro Thr Phe
85 90 95

Gly Arg Trp Asp Ala Val Leu Met Asp Gln Leu Asp His Arg Ala Ser
100 105 110

Tyr Phe Pro Val Gly Val Gly His Ala Phe Val Ala Leu Glu Asp Asp
115 120 125

Thr Asp Met Ser Tyr Met Leu Ser Gly Arg Tyr Val Ala Glu His Glu
130 135 140

Leu Ser Leu Ser Ala Leu Asp Pro Asp Leu Gly Leu Pro Ile Pro Thr
145 150 155 160

Asp Leu Glu Pro Ile Leu Ser Glu Arg Asp Arg Ala Ala Val Thr Leu
165 170 175

Ala Glu Ala Gln Glu Lys Gly Leu Leu Pro Asp Tyr Ala Arg Cys Gln
180 185 190

Glu Ile Glu Arg Gly Leu Val Pro Arg Ala Arg Pro Ala Ala
195 200 205

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 6085 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:
(A) NAME/KEY: misc_RNA
(B) LOCATION:378..1665
(D) OTHER INFORMATION:/function= "ORF"

(ix) FEATURE:
(A) NAME/KEY: misc_RNA
(B) LOCATION:1747..2553
(D) OTHER INFORMATION:/function= "ORF"

(ix) FEATURE:

(A) NAME/KEY: misc_RNA

(B) LOCATION:2593..4011

(D) OTHER INFORMATION:/function= "ORF"

(ix) FEATURE:

(A) NAME/KEY: misc_RNA

(B) LOCATION:4013..4999

(D) OTHER INFORMATION:/function= "ORF"

(ix) FEATURE:

(A) NAME/KEY: misc_RNA

(B) LOCATION:5071..6085

(D) OTHER INFORMATION:/function= "ORF"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

TGATGGCCCA GCACTTCGGC GAGTGCCCGG ACGCCAGTCT GCGGCGGTCG GACCTGATGA 60

ACGGGGCGAT CGACATGATG ACGGGCCTGA TGCGGCCGCT GGCGGAACTG CTCGTCACCC 120

TGCCGTCGGG GCGCCGGGGC ATGACCGCCG GACCGTCCTT CGAACTGCCC GAGCAGCCCG 180

CGCCCGTGTC CCGGCCGGAC GTGGCCAGAC GCGGTATCGC CCGCCGCCTC GACGACCTCG 240

CGGCGCAGTG CGCCAAGCAT CCGCTCGTCC CCCCGCGCGT GGCGGAGATG AGCACCTTCT 300

GGGCCGACCG CTTCCGCCCG CCGAGCCGTT AGGGCCGGTT GCGAAAGGGG CCGAACACTT 360

CCGACCGAAG GAGACGCATG CCATCCGCGA CGCTGCCGCG GTTCGACCTC ATGGGCTGGG 420

ACAAGGAGGA CATCGCCCAC CCCTACCCGG TCTACCGGCG CTACCGGGAG GCCGCCCCGG 480

TCCATCGCAC GGCGGCGGGC CCCGGAAAGC CTGACACCTA CTACGTGTTC ACCTACGACG 540

ACGTGGTCCG CGTCCTGTCC AACCGGCGGT TCGGCCGCAA CGCCCGCGTG GCCTCCGGCG 600

ACACCGGCCC CGACACCGCG CCCGTCCCGA TCCCGGCCGA GCACCGCGCC CTGCGGACCG 660

TCGTCGAGAA CTGGCTGGTC TTCCTCGACC CCCCGCGCCA CACCGAACTG CGCTCCCTGC 720

TCACCGGCGA GTTCTCACCC TCGATCGTCA CCGGCCTGCG CCCCCGCATC GCCGAACTCG 780

CGAGCGAACT CCTGGACCGG CTCCGAGCAC ACCGCCGGCC CGATCTCGTC GAGGGTTCGC 840

GGCGCCCCTC CCCGATCCTC GTCATCTCCG CACTGCTGGG CATCCCCCGC GGAGGACCAC 900

ACCTGGTGCG CGCCAACGCG GTGGCCCTTC AGGAGGCCGG CACCACGCTC GCGCGGCGGC 960

CACGGTACGC ACGGGCCGAG GCGGCGTCCC AGGAGTTCAC CCGCTACTTC CGGCGAGAGG 1020

TGGACCGGCG CGGCGGCGAC GACCGCGACG ATCTGCTCAC CCTCCTCGTC CGCGCCCGGG 1080

ACACCGGATC ACCGCTCAGC GTGGACGGCA TCGTCGGCAC CTGCGTCCAT CTGCTCACCG 1140

CCGGCCACGA GACCACCACC AACTGCCTCG CCAGGGCGGT CCTCACCCTG CGCGCCCACC 1200

CTGACGTCCT CGACGAGCTG CGCACCACAC CGGAGTCGAC ACCGGCGGCC GTCGAAGAGC 1260

TGATGCGGTA CGACCCGCCC GTGCAGGCGG TGACGCGCTG GGCGTACGAG GACATCCGGC 1320

TCGGCGACCA CGACATCCCG CGCGGCAGCC GGGTGGTCGC GCTGCTGGGC TCGGCGAACC 1380

GGGACCCGGC GCGCTTCCCG GCTCCCGACG TGCTGGACGT CCACCGCGCC GCCGAACGGC 1440

AGGTGGGCTT CGGCCTCGGA ATCCACTACT GCCTCGGCGC GACCCTGGCC CGCGCCGAGG 1500

CCGAGATCGG TCTGAGGGCC CTGCTGGACG GCATCCCCGC CCTCGGCCGA GGCGCCCACG 1560

AGGTCGAGTA CGCCGACGAC ATGGTCTTCC ACGGCCCGAM GCGGCTCCTC CTCGACCTGC 1620

CGGAMGCCAC GTDCCCCTCG GCCAGCCACC CCTAGCCCTC GGCCACCCCT CGACCCCGGC 1680

CATCCCTTGC CCTGGCCACC CCTCGACCCC GGCCCTCTCG ACTCGCACCA GCAGGAAGGC 1740

ACATCCATGA CGCAGCAGTC CGACACCACC GCCGACTCGG TCGGTGAGGT GTACGACCAG 1800

TTCGCCGACG CCGGCGCCAG CACCGCGATG GGCGGCAACA TCCACGTGGG GTACTGGGAC 1860

GACGACCCCG AGGTGCCGAT CGCCGAGGCC ACCGACCGGC TCACCGATCT CGTCGCCGAG 1920

CGCCTCGCGC TCCGCCCCGA CCGGCATCTG CTGGACGTGG GCTGCGGCAT CGGCGTGCCG 1980

GCTCTCAGGA TCGCCGGAGC GCACGACGTC CGCGTCACCG GGATCACCGT CAGCCAGCAG 2040

CAGGTCACCG AGGCGGCCGA GCGGGCGGTG GAGTCCGATG CCGGGGGCCG GGTCTCCTTC 2100

CGGCTGGCGG ACGCCATGGA CCTCCCCTTC GAGGACGTCT CCTTCGACGG CGCCTTCGCC 2160

ATCGAGTCGC TGCTGCATCT GCCCGACCAG ACACCCGCGC TCAAGGAGAT CCACCGGGTC 2220

GTCCGCCCCG GCGGCCGGCT CGTCATCGCC GACCTGTGTC AGCGACAGCC GTTCACCGGC 2280

GCCGACAAGG AGGTGCTCGA CGGGATGCTG CTGATGTACG AGATCGCCGG GATCAACACA 2340

CCCTACGAGC ATCGCGCGCG ACTGGCGGAG GCGGGCTGGG AACTGCTGGA GCTGACGGAC 2400

ATCGGTGAGC AGGTCCGCGC CTACTACGGG CATGCCGCCG CCGCGTTCCG GGGTCTCGCC 2460

GGGGCTCTCG ACGCCGGCGC GGCGCAGCAG ATGAACGCGG CGGCCGACCT GATGGAGGCT 2520

TCGGAGGGCA TCCGCACTCC GGTTACGTCC TGATCACGCG CAGCGGTCCT GACCGGACGG 2580

GGAGACCTGT GATGTCTTCT GGTCTCGGCC CGCCGTCCGC CGCCGTACGC CCGCGTGAGG 2640

ACCGTGCGAC GGCCGACCGT GTCGCCCTGT CCGCCGCGAC CGCCCGCGGA GCACCGGTCG 2700

CGGACCGAGG AGGTGCGGGC CTGGCTGGCC GAGCGGCGCC GGGCCCATGT GTTCGAGGTG 2760

ACGCGGATAC CGTTCGCGGA GCTGCGGCAG TGGCGGTTCG AGGAGGGCAC CGGCAATCTC 2820

GTGCACCGCA GCGGACGGTT CTTCACCGTC GAGGGCATGC ATGTCGTCGA GTCGGACGGG 2880

CCCTTCGGGG ACGGCCCGTA CCAGGAGTGG CAGCAGCCCG TCGTCCGCCA GCCCGAAGTG 2940

GGCATCCTCG GCATTCTCGC CAAGGAGTTC GACGGAGTGC TGCACTTCCT GATGCAGGCC 3000

AAGATGGAGC CGGGCAATCC CCGTCTGCTC CAGCTCTCCC CGACCGTGCA GGCCACCCGC 3060

AGCAACTACA CACGGGCTCA CCGGGGCACG GACGTCAAGC TCATCGACCA TTTCTTCCGA 3120

CCCGACCCCG ACCGGGTCCT CGTCGACGTC CTGCAGTCCG AACAGGGCTC GTGGTTCTAC 3180

CGCAAGTCCA ATCGCAACAT GATCGTGGAG ACCGTCGACG ACGTTCCCGA ACTGGACGAC 3240

TTCCGCTGGC TCACCCTCGG CCAGATCGCC GAACTGCTGC ACGAGGACGA CCTGGTCAAC 3300

ATGAACGCCA GGACGGTGCT GTCGTGCGTG CAGTACCCCG ACACCTCGCC CGGGGCGCTG 3360

CTCTCCGACG CCCAGCTCCT GTCCTGGTTC ACCGGGGAGC GTTCCCGGCA CGACATCCGC 3420

GTGGAGGCGG TGCCGCTCGC TCCGTGCGCG GCCTGGAAGC AGGGTGTCGA GGCGATCGAG 3480

CACGAGAACG GGCGCTACTT CAAGGTCGTC GCCGTCTCCG TGCGGGCCGG CAACCGCGAG 3540

GTGGTCGACT GGGACCAGCC GTTGCTGGAG CCGGTGGGCC TGGGGGTCAG CGCCTTCCTG 3600

GTGCGCGAGA TCGAGGGCGT ACCCCATGTC CTGGTCCATG CCCAGGCCGA GGGCGGGTTC 3660

CTGGACACCG TCGAGCTGGC TCCGACCGTC CAGTGCACAC CCGGCAATTA CGCCCATCTC 3720

ACCCCGGAGC ACCGCCCGCC GTTCCTCGAC ACCGTCCTCG ACGCCCGCCC CGAGCGCATC 3780

CGCTACGAGG CCGTCCACTC CGAGGAGGGC GGACGCTTCC TCAACGCCAG GAGCCGCTAT 3840

CTGCTGGTCG ACGCCGACGA CGTCCCCCTC GCCCCGCCCC CCGGCTACAC CTGGGCCACC 3900

CCGGGCCAGC TCAGGACCCT CACCCGGCAC GGCCACTACC TGAACGTCGA GGCCCGCACG 3960

CTGCTGGCCT GCGTCAACGC GACGGCCGCA GGGCCGCGAG GAGGACAGTG ACATGGGGAA 4020

CCCACCGCTG ATCACCGTGC TCGGTGCCTC GGGTTTCGTC GGGTCGGCCG TCACCCGGGC 4080

GCTGGCGTCC CGGCCCGTCC GGCTCCGGCT CGTCTCCCGT CGGCCCTGCG TCCCCTCCCC 4140

CGGCCCGGCC GAGACCGATG TCGTCACCGC CGATCTCACC GACCGGGCCG CGCTGGCCGG 4200

GGCGGTGCAG GGTTCGGACG GGGTGATCCA TCTGCTGCTG GGGGAGGGCG GCTGGCGGGC 4260

AGCCGAGTCC GACCCCGGTG CCGAGCACGT CAACGTCGGC GTCATGCGGG ACCTCGTCGA 4320

GGTACTGCGG CCCGCGCCCG GCGACGCGGC ACCCCCGCTG GTGGTGTACG CCGGTGCCGC 4380

CTCGCAGGTC GGGGTGCCGC CCCGGGAGCC CCTCGACGGC AGCGAGCCCG ACCGCCCGGA 4440

GACCGCCTAC GACCGGCAGA AACTGACCGC TGAACACCTC CTGCTCAAGG CCACCGCCGA 4500

GGGCCGGGTA CGCGGCATCG GCCTGCGTCT GCCCACCGTG TTCGGCGAGA GCACGGCGTC 4560

CGGCACCGGC GACCGAGGCG TCGTGTCGGC CATGGCGCGC AAGGCCCTCG ACGGGCAGAC 4620

GCTCACCATG TGGCACGACG GCACCGTGCG CCGCGACCTG GTCCATGTCG ACGATGTCGC 4680

GGCGGCGTTC ACGGCCGCCC TCGACCACCC GGACGCCCTC GTGGGCGGCM ATTGGCTGAT 4740

CGGGGCCGGC CGGGGCGACG CGCTCGGCGA TGTCTTCCGG CTGATCGCCC TCACCGCGGC 4800

CGATGTCCTC GGGCGGTCCC CGGTCGACGT GGTCTCCGTA GAACCGCCCG CGCACGCCCC 4860

CGTGACCGAC TTCCGCAGCG TCACCCTCGA CTCCTCGCGT TCCGCGCGGC CACCGGTTGG 4920

CGCCCCCGGA ATCTCCCTGC CCGAGGGCGT GCGCCGCACC GTCACCGCCC TGGCCCGGGA 4980

GCGGGCCGCG AGCCGGTGAC GTCAGCGCCC CCGACCCCTA CTCACCACAG GCGTACGGCC 5040

GTGCGCCCGC AGTACTGGAA AGGCTGGACG ATGACCACGC GTGTATGGGA CTACCTGGCG 5100

GAGTACCGAG CCGAGCGGGC GGACATCCTC GACGCCGTCG AAACGGTCTT CGAGTCGGGC 5160

CAGTTGGTGC TCGGCGCGAG TGTGCGCGGC TTCGAGGAGG AGTTCGCCGC ATACCACGGA 5220

GTGGACCACT GCGTGGGTGT CGACAACGGA ACGAACGCCA TCAAGCTCGC TCTCCAGGCC 5280

CTCGGGGTCG GCCCCGGCGA CGAGGTGATC ACGGTGTCCA ACACCGCCGC CCCCACCGTC 5340

GTCGCCATCG ACTCCACCGG CGCCACCCCC GTCTTCGTCG ACGTCCGCGA GGACGACTTC 5400

CTCATGGACA CGAGCCAGGT CGAGGCGGCC GTCACCGAAC GCACCCGCTG CCTGCTCCCG 5460

GTCCACCTGT ACGGCCAGTG CGTCGACATG GCGCCGCTGA AGGAGATCGC CGCCCGGCAC 5520

GTGGTCGTCC TGGAGGACTG CGCCCAGGCC CATGGCCGAC AGGGCGACAC CATGGCCGGC 5580

ACCACCGGTG ACGCCGCCGC CTTCTCCTTC TACCCGACCA AGGTCCTCGG CGCGTACGGC 5640

GACGGCGGCG CCACGATCAC CGGCGACGCG TCCGTGGCCG CCCGCCTGCG ACGCCTGCGC 5700

TACTACGGCA TGGACGAGCG CTACTACACC CTGGAGACCC CCGCCCACAA CAGCCGCCTG 5760

GACGAACTCC ACGCAGAGAT CCTCCGCCGC AAACTTCGGC GCCTCGACAC CTACGTCAAG 5820

GGCCGCCGCG CCGTCGCCGA ACGCTACGCC GACGGGCTCG CCGACACCGA CCTCGTCCTG 5880

CCGCACACGG TCCCCGGCAA CGAGCACGTC TACTACGTGT ACGTCGTCCG CCACCCCCGG 5940

CGTGACGACA TCATCGAGCG CCTCAAGGCC CACGACGTCC ACCTCAACAT CAGCTATCCG 6000

TGGCCGGTGC ACACCATGAC GGGCTTCGCC CACCTCGGCT ACGCAAGGGC TCGCTCCCGG 6060

TCACCGAGGC ACTGGCGCGA GATCT 6085

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1845 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

AGATCTACCG CTACCGGCGA GGCACGGCCG CGAACCGCCC GGCCTTCGTC CATACCCCCG 60

AGCCCGATCA GATCTGCCCC GCCCACTGGC TCAACCCGGT GCTGATCGAG GCCGTGGGCG 120

TCCACCCGGA CGGCCCCCTG CTACTGAGTA CGACCGTCGA CGGCGTGGTC CAGACCGACG 180

ACCACGTCGA GGCCACCCTC ACCGACCACG CCACCGGCAC CACCGGCACC GTCCGGGCAC 240

GCTTCCTCGT CGCCTGTGAC GGCGCCTCCT CGCCCGTCCG CCGGCGCTGC GGCATCGAGG 300

CACCGGCCCG CCACCGTACG CAGGTCTTCC GCAACATCCT CTTCCGCGCC CCCGAGCTCA 360

AGGACCGCCT GGGCGAGCGG GCCGCCCTGG TCCACTTCCT GATGCTGTCG TCCACCGTGC 420

GCTTCCCCCT GCGCTCGCTG AACGGCAGCG ACCTGTACAA CCTGGTCGTG GGCGCCGACG 480

ACGACACCGG CGCCCGACCC GACGTCCCTG GCCCTGCAGT GATCAAGGAC GCCCTGGCCC 540

TCGACACCCC GGTGGAGCTG CTCGGCGACA GCGCGTGGCG TCTCACCCAC CGTGTCGCCG 600

ACCGCTACCG GGCCGGACGG ATCTTCCTCG CCGGCGACGC CGCGCACACC CTGTCGCCCT 660

CCGGCGGCTT CGGCCTCAAC ACCGGTATCG GCGACGCCGC CGATCTCGGC TGGAAGCTCG 720

CCGCCACCCT GGACGGCTGG GCCGGGCGGC ACCTCCTCGA CACCTACGAC AGCGAGCGTC 780

GACCGATCGC CGAGGAGAGC CTGAACGAGG CCCACGACAA TCTTCGGCGC ACCATGAAAC 840

GGGAGGTCCC GCCGGAGATC CACCTCGACG GACCCGAGGG CGAGCGGGCC CGCGCCGTGA 900

TGGCCAGGCG CCTCGAGAAC AGCGGCGCGC GGCGGGAGTT CGACGCCCCG CAGATCCACT 960

TCGGACTGCG CTACCGCTCC TCGGCGATCG TCGACGACCC CGACGTACCG GTCCGCCAGG 1020

GGCAGCCGGA CGCCGATTGG CGGCCCGGCA GCGAGCCCGG GTACCGCGCC GCGCACGCCT 1080

GGTGGGACTC CACGACCTCC ACGCTCGACC TCTTCGGCCG CGGCTTCGTC CTGCTCCGCT 1140

TCGCGGACCA CGACGGCCTC CCGGCGATCG AGCGCGCGTT CGCCGAGCGG GGCGTACCCC 1200

TGACCGTGCA CCAGGGACAC GACACGGAGA TCGCCAAGCT GTACGCACGC TCCTTCGTCC 1260

TGGTCCGCCC CGACGGTCAT GTCGCCTGGC GCGGCGACGA CCTGCCCGGC GACCCGACGG 1320

CCCTGGTCGA CACGGTGCGG GGTGAGGCCG CGCCCCGTGA ACCGCGGGGC TGAGGCCCAC 1380

GCGGCCTCCC GTCCGCCGAT GGGGCGGCTC GGACCGAAGC TCCTCTGACC TGTATGTTCC 1440

CACAGTCCGT GCACGGTGCG GACCCTGTAG GGACGCCCGG TAAACTCCGT ACACGTGACT 1500

TCTGCGCCAG CCAAGCCCCG CATCCCGAAC GTCCTCGCCG GACGCTACGC CTCCGCCGAG 1560

CTCGCCACGC TCTGGTCCCC CGAGCAGAAG GTGAGGCTGG AGCGGCAGCT CTGGCTGGCC 1620

GTGCTGCGGG CCCAGAAGGA CCTCGGCATC GAGGTGCCGG ACGAGGCGCT CGCCGACTAC 1680

GAGCGGGTCC TCGACACCGT CGACCTGGCC TCCATCGCCG AGCGCGAGAA GGTCACGCGG 1740

CACGACGTGA AGGCGCGGAT CGAGGAGTTC AACGACCTCG CCGGGCACGA GCACGTGCAC 1800

AAGGGCATGA CCTCCCGGGA CCTCACGGAG AACGTCGAGC AGCTG 1845

What is claimed is:

1. An isolated DNA fragment comprising a DNA region that is involved directly or indirectly in the biosynthesis of indole-carbazole alkaloids, including the adjacent DNA regions to the left and right which, because of their function in connection with indole-carbazole alkaloid biosynthesis, qualify as constituents of the indole-carbazole alkaloid gene cluster; and functional fragments thereof.

2. A DNA fragment according to claim 1, wherein the indole-carbazole alkaloid is staurosporin.

3. A DNA fragment according to claim 1, which comprises a DNA region that is involved directly or indirectly in the biosynthesis of staurosporin.

4. A DNA fragment according to claim 1, wherein the said DNA region is obtainable from the gene cluster within the genome of *Streptomyces longisporoflavus* that is responsible for staurosporin biosynthesis.

5. A DNA fragment according to claim 1, which fragment comprises a 35 kb DNA region (Figure 2).

6. A DNA fragment according to claim 1, which fragment comprises a 10 kb region (Figure 1).

7. A DNA fragment according to claim 1, which fragment contains one or more of the partial nucleotide sequences set out in SEQ ID NOs 1, 4 and 5, or functional fragments thereof, and any further DNA sequences in the vicinity of that sequence that, on the basis of homologies present, can be regarded as structural or functional equivalents and are therefore capable of hybridising with that sequence.

8. A DNA fragment according to claim 1, which fragment contains the partial nucleotide sequence set out in SEQ ID NO 1, 4 or 5.

9. A DNA fragment according to claim 1, wherein 2 or 1 DNA fragment(s) of the following size, which are obtainable by the method according to the invention from the *Strepto-*

myces longisporoflavus genome, overlap(s) with the 2.1 kb fragment according to Figure 2: EcoRI: > 20 kb, PvuII: 3.5 kb and 6.5 kb; PvuI: 3.4 kb and 2.1 kb; BclI: 3.6 kb.

10. A DNA fragment according to claim 1, which fragment contains portions of sequence having homologies to enzymes that are involved in the synthesis of indole-carbazole alkaloids.

11. A DNA fragment according to claim 1, which fragment contains portions of sequence having homologies to the methyl transferases and the amino transferases of *Streptomyces* or *Actinomyces* or to dTDP-4-keto-6-deoxyglucose 3,5-epimerases.

12. A DNA fragment according to claim 1, which fragment contains portions of sequence that code for a methyl transferase.

13. A DNA fragment according to claim 1, which fragment contains portions of sequence having homologies to the 35 kb DNA region according to claim 5, to the10 kb DNA region according to claim 6, or to SEQ ID NOs 1, 4 or 5 according to claim 7 and can therefore be used as a hybridisation probe within the genomic gene bank of an indole-carbazole alkaloid-producing organism for detecting constituents of the gene cluster responsible therefor.

19. A DNA fragment according to claim 1, which DNA fragment comprises exclusively genomic DNA.

20. A DNA fragment according to claim 1, which fragment contains the partial nucleotide sequence set out in SEQ ID NO 1, 4 or 5 or a sequence that, on the basis of homologies present, can be regarded as a structural or functional equivalent of the said partial sequence and is therefore capable of hybridising with that sequence.

21. A DNA fragment according to claim 1, which codes for the protein set out in SEQ ID NO 2 or SEQ ID NO 3, for the proteins represented by the open reading frames in SEQ ID NO 4, or for a functional derivative thereof.

22. A hybrid vector containing a DNA fragment according to claim 1.

23. A hybrid vector containing an expression cassette containing a DNA fragment according to claim 1.

24. A host organism containing a hybrid vector according to claim 22.

25. A host organism into the chromosome of which a DNA fragment according to claim 1 has been integrated.

26. A method of identifying, isolating and cloning a DNA fragment that is obtainable from the gene cluster within the genome of *Streptomyces* or *Actinomyces* that is responsible for staurosporin biosynthesis and that contains at least one gene that is involved directly or indirectly in the biosynthesis of indole-carbazole alkaloids; which method comprises the following steps:

a) constructing a representative gene library of an indole-carbazole alkaloid-producing organism from the group of the *Streptomyces* or *Actinomyces*, which library contains substantially the entire genome divided into individual clones,
b) screening the said clones using a specific DNA probe that hybridises at least with a portion of the gene cluster responsible for the indole-carbazole alkaloid biosynthesis,
c) selecting the clones that allow a hybridisation signal with the DNA probe to be recognised; and
d) isolating a DNA fragment from the said clone that contains at least one gene that is involved directly or indirectly in the biosynthesis of the indole-carbazole alkaloid.

27. A method according to claim 26, wherein the said staurosporin-producing organism is *Streptomyces longisporoflavus*.

28. A method according to claim 26, wherein the said hybridisation probe is a DNA fragment according to claim 1.

29. A method according to claim 26, wherein there are used as hybridisation probe sections of sequence originating from the right- and/or left-hand margins of the said DNA fragments.

30. A method according to claim 26 of identifying and isolating all of the DNA sequences that are involved in the indole-carbazole alkaloid gene cluster, which method comprises

a) constructing a representative gene library of an indole-carbazole alkaloid-producing organism from the group of the *Streptomyces* or *Actinomyces*, which library contains subtantially the entire bacterial genome divided into individual clones;
b) hybridising the said clones using as probe molecule one of the previously isolated DNA fragments or selected portions thereof that overlap at least with a portion of the adjacent DNA regions to the right and/or left within the gene cluster;
c) selecting the clones that allow a strong hybridisation signal with the DNA probe to be recognised;
d) isolating the fragments that contain overlapping DNA regions from the clones selected in accordance with (c) and isolating the fragment that projects furthest beyond the overlapping region;
e) testing the DNA fragment isolated in accordance with (d) for its ability to function within the gene cluster;
f) if it can be demonstrated that the DNA fragment isolated in accordance with (d) functions in the context of indole-carbazole alkaloid biosynthesis, repeating the method according to steps (a) to (e), the DNA fragment isolated in accordance with (d), or selected portions thereof, especially those from the left- and/or right-hand margin of the said fragment, now acting as the DNA probe, until in the function test for each newly isolated DNA fragment no further functioning is detected in the indole-carbazole alkaloid biosynthesis and the end of the gene cluster has thus been reached; and
g) carrying out the method according to steps (a) to (f), if necessary in the other, not hitherto selected, direction.

31. A method according to claim 30, wherein the said organism is *Streptomyces longisporo-flavus.*

32. The use of DNA fragments according to claim 1 in the preparation of indole-carbazole alkaloids, indole-carbazole alkaloid derivatives or precursors.

33. The use of DNA fragments according to claim 1 for inactivating genes of the indole-carbazole alkaloid biosynthesis.

34. The use of DNA fragments according to claim 1 in PCR amplification.

35. The use of DNA fragments according to claim 1 in the preparation of indole-carbazole alkaloids, indole-carbazole alkaloid derivatives or precursors.

36. The use of a hybrid vector according to claim 22 in the preparation of indole-carbazole alkaloids, indole-carbazole alkaloid derivatives or precursors.

37. The use of a hybrid vector according to claim 22 in the preparation of staurosporin, staurosporin derivatives or precursors.

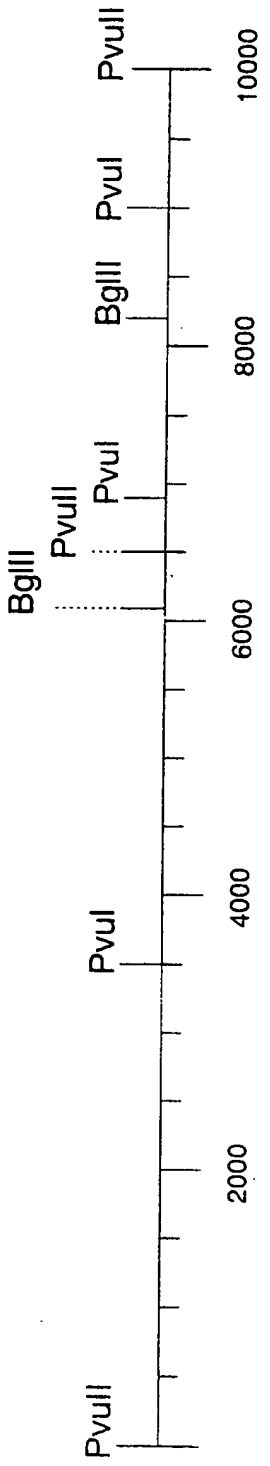

PvuII
PvuI
BglII
PvuI
PvuII
BglII
PvuI
PvuII
10000
8000
6000
4000
2000

Fig. 1

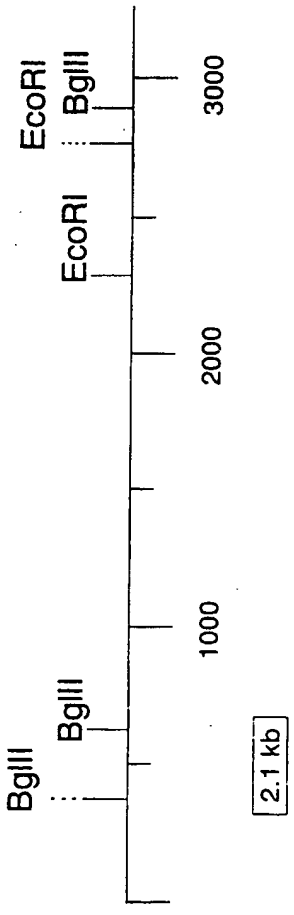


Fig. 2

INTERNATIONAL SEARCH REPORT

Inte onal Application No
PCT/EP 96/03643

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/52 C12N15/10 C12N9/92 C12N9/10 C12N1/21
C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO,A,95 00520 (CIBA GEIGY AG ;HOEHN THIERRY PASCALE (FR); GHISALBA ORESTE (CH); P) 5 January 1995 see the whole document ---	1-37
A	EP,A,0 444 503 (SQUIBB BRISTOL MYERS CO) 4 September 1991 see the whole document ---	1-37
A	JOURNAL OF NATURAL PRODUCTS, vol. 51, no. 5, 1 January 1988, pages 893-899, XP000561179 MEKSURIYEN D ET AL: "BIOSYNTHESIS OF STAUROSPORINE, 2. INCORPORATION OF TRYPTOPHAN1,2" see the whole document ---	1-37
	-/--	

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
3 December 1996	06.12.96

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+ 31-70) 340-2040, Tx. 31 651 epo nl, Fax (+ 31-70) 340-3016	Hornig, H

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	MOLECULAR & GENERAL GENETICS, vol. 241, no. 1-2, October 1993, SPRINGER INTERNATIONAL, AMSTERDAM, NL, pages 193-202, XP002020055 H. KRÜGEL ET AL.: "Nucleotide sequence analysis of five putative Streptomyces griseus genes, one of which complements an early function in daunorubicin biosynthesis that is linked to a putative gene cluster involved in TDP-daunosomine formation" see the whole document ---	1-37
A	TRENDS IN GENETICS, vol. 11, no. 6, June 1995, ELSEVIER SCIENCE LTD., AMSTERDAM, NL, pages 217-218, XP002020056 A. POSPIECH AND B. NEUMANN: "A versatile quick-prep of genomic DNA from gram-positive bacteria" cited in the application see the whole document ---	1-37
A	US,A,4 973 552 (SCHROEDER DANIEL R ET AL) 27 November 1990 see the whole document ---	1-37
A	J. ANTIBIOT. (1995), 48(5), 428-30 CODEN: JANTAJ;ISSN: 0021-8820, May 1995, XP002020057 GOEKE, K. ET AL: "Production of the staurosporine aglycon K-252c with a blocked mutant of the staurosporine producer strain Streptomyces longisporoflavus and by biotransformation of staurosporine with Streptomyces mediocidicus ATCC 13279" see the whole document ---	1-37
A	J. ANTIBIOT. (1995), 48(4), 300-5 CODEN: JANTAJ;ISSN: 0021-8820, April 1995, XP002020058 HOEHN, PASCALE ET AL: "3'-Demethoxy-3'-hydroxystaurosporine, a novel staurosporine analog produced by a blocked mutant" cited in the application see the whole document ---	1-37

-/--

1

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	J. ANTIBIOT. (1995), 48(2), 143-8 CODEN: JANTAJ;ISSN: 0021-8820, February 1995, XP002020059 CAI, YANG ET AL: "A nitro analog of staurosporine and other minor metabolites produced by a Streptomyces longisporoflavus strain" see the whole document -----	1-37

INTERNATIONAL SEARCH REPORT

Information on patent family members

Int ional Application No
PCT/EP 96/03643

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
WO-A-9500520	05-01-95	AU-A-	7000094	17-01-95
		EP-A-	0703917	03-04-96
EP-A-0444503	04-09-91	US-A-	4973552	27-11-90
		CA-A-	2036669	21-08-91
		JP-A-	3244387	31-10-91
		JP-B-	7000037	11-01-95
US-A-4973552	27-11-90	CA-A-	2036669	21-08-91
		EP-A-	0444503	04-09-91
		JP-A-	3244387	31-10-91
		JP-B-	7000037	11-01-95